

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,

please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB#

54203

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

BEST AVAILABLE COPY

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher <u>P. Schreiber</u>	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone # <u>308-4272</u>	AA Sequence (#) _____	Dialog _____
Searcher Location <u>CM 112E18</u>	Structure (#) _____	Questel Orbit _____
Date Searcher Provided <u>11/12</u>	Bibliographic _____	Dr. Lina _____
Date Completed <u>11/15</u>	Citation _____	Lexis Nexis _____
Searcher Prep & Review Time <u>5</u>	Fulltext _____	Sequence Systems <u>CompuGel</u>
Client Prep Time _____	Patent Family _____	WPAW Internet _____
Printing Time <u>10</u>	Other _____	Other Specified _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:47:29 ; Search time 10414.5 Seconds
(without alignments)
12052.456 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatgtacggcgccagatatac.....gcgtgtcgtctagcagcgttc 8115

Scoring table:

IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rod:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vi1:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rod1:*
95: gb_rod2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3535	43.6	3660	3 LBAGL	M23530 L.delbrueck
2	3497.2	43.1	5059	9 AX009488	AX009488 Sequence
3	3386.6	41.7	5015	3 LBALAC2BUL	M55068 L.delbrueck
4	1336.4	16.5	3840	56 EVEI32038	AJ132038 Expressio
5	1240.4	15.3	3853	9 AR098190	AR098190 Sequence
6	1238.4	15.3	2093	3 LP1PLPREP	M31223 Plasmid pLP
7	1193.8	14.7	2140	2 C300RFR	J03319 Plasmid pC3
8	1082	13.3	3986	56 PCDNA32EO	X90639 Cloning vec

C_35	906.4	11.2	3024	3	PK9CCDB
C_44	906.4	11.2	2999	56	SYNCCDB
C_43	906.4	11.2	2999	56	SYNCCDB
C_42	906.4	11.2	2860	56	AF017063
C_41	906.4	11.2	2820	56	AF022336
C_40	906.4	11.2	2769	56	AX018504
C_39	906.4	11.2	2684	56	AF016541
C_38	906.4	11.2	2661	56	SYNKMACG
C_37	906.4	11.2	2000	10	I03343
C_36	906.8	11.2	7511	56	CVU37068
C_35	906.8	11.2	7511	56	CVU37023
C_34	907	11.2	3141	56	AF334726
C_33	907	11.2	3018	56	AF334725
C_32	908.2	11.2	4109	56	PSPORT1
C_31	909.2	11.2	7758	56	AF063849
C_30	909.2	11.2	6706	56	AF063850
C_29	909.2	11.2	6631	56	AF063585
C_28	909.2	11.2	6525	56	AF063848
C_27	921.8	11.4	6285	10	I722638
C_26	921.8	11.4	6285	9	AR066551
C_25	921.8	11.4	6285	9	AR066551
C_24	921.8	11.4	6285	9	AR018923
C_23	928.6	11.4	6094	56	AX027912
C_22	928.6	11.4	6094	56	AX027906
C_21	935.2	11.5	5729	56	AX027914
C_20	935.2	11.5	5729	9	AX027908
C_19	935.6	11.5	5432	9	AX026821
C_18	959	11.8	5732	56	AX027910
C_17	959	11.8	5732	56	AX027904
C_16	968.2	11.9	4657	56	AF053406
C_15	968.6	11.9	4627	56	AF053409
C_14	968.6	11.9	4627	56	AF053408
C_13	968.6	11.9	4053	56	AF053407
C_12	973	12.0	4283	10	I63120
C_11	988	12.2	4587	9	AX060344
C_10	988	12.2	4587	9	AR098191
C_9	1068.6	13.2	4026	9	AR098191
C_8	906.4	11.2	3024	3	PK9CCDB
C_7	906.4	11.2	2999	56	SYNCCDB
C_6	906.4	11.2	2860	56	AF017063
C_5	906.4	11.2	2820	56	AF022336
C_4	906.4	11.2	2769	56	AX018504
C_3	906.4	11.2	2684	56	AF016541
C_2	906.4	11.2	2661	56	SYNKMACG
C_1	906.4	11.2	2000	10	I03343
C_0	906.4	11.2	7511	56	CVU37068
C_-1	906.4	11.2	7511	56	CVU37023
C_-2	907	11.2	3141	56	AF334726
C_-3	907	11.2	3018	56	AF334725
C_-4	908.2	11.2	4109	56	PSPORT1
C_-5	909.2	11.2	7758	56	AF063849
C_-6	909.2	11.2	6706	56	AF063850
C_-7	909.2	11.2	6631	56	AF063585
C_-8	909.2	11.2	6525	56	AF063848
C_-9	921.8	11.4	6285	10	I722638
C_-10	921.8	11.4	6285	9	AR066551
C_-11	921.8	11.4	6285	9	AR018923
C_-12	928.6	11.4	6094	56	AX027912
C_-13	928.6	11.4	6094	56	AX027906
C_-14	935.2	11.5	5729	56	AX027914
C_-15	935.2	11.5	5729	9	AX027908
C_-16	935.6	11.5	5432	9	AX026821
C_-17	959	11.8	5732	56	AX027910
C_-18	959	11.8	5732	56	AX027904
C_-19	968.2	11.9	4657	56	AF053406
C_-20	968.6	11.9	4627	56	AF053409
C_-21	968.6	11.9	4627	56	AF053408
C_-22	973	12.0	4283	10	I63120
C_-23	988	12.2	4587	9	AX060344
C_-24	988	12.2	4587	9	AR098191
C_-25	1068.6	13.2	4026	9	AR098191

ALIGNMENTS

AR0908191 Sequence
 AX060304 Sequence
 I6120 Sequence 3
 I8549 Sequence 1
 AF053407 Expressio
 AF053408 Expressio
 AF053409 Expressio
 AF053406 Expressio
 AX027904 Sequence
 AX027910 Sequence
 AX026821 Sequence
 AX027988 Sequence
 AX027914 Sequence
 AX027906 Sequence
 AX027912 Sequence
 AR018923 Sequence
 AR066551 Sequence
 I72268 Sequence 49
 I76833 Sequence 49
 AF063848 Cloning v
 AF063855 Cloning v
 AF063850 Cloning v
 AF063849 Cloning v
 U12390 Cloning vec
 AE524725 Cloning v
 AE524726 Cloning v
 U37023 Cloning vec
 U37068 Cloning vec
 I03343 Sequence 3
 M17626 Synthetic P
 AF016541 Cloning v
 AX018504 Sequence
 AF022336 Cloning v
 AF017623 Cloning v
 L38498 Cloning vec
 L38499 Cloning vec
 L27082 Plasmid K1L

```

RESULT 1
LBRAGL
LOCUS          LBAGAL          3660 bp      DNA          BCT          26-APR-1993
DEFINITION    L.delbrueckii bulgaricus beta-galactosidase gene; complete cds.
ACCESSION     M23530
VERSION       M23530.1  GI:149546
KEYWORDS      galactosidase.
SOURCE        L.delbrueckii bulgaricus (strain B131) DNA.
ORGANISM      Lactobacillus delbrueckii
              Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
              Lactobacillus.
REFERENCE     1 (bases 1 to 3660)
AUTHORS      Schmidt,B.F., Adams,R.M., Reguadt,C., Power,S. and Mainzer,S.E.
TITLE         Expression and nucleotide sequence of the Lactobacillus bulgaricus
              beta-galactosidase gene cloned in Escherichia coli
JOURNAL       J. Bacteriol. 171, 625-635 (1989)
MEDLINE       89123132
COMMENT       Draft entry and computer-readable sequence for [1] kindly submitted
              by B.F.Schmidt, 29-MAR-1989.
FEATURES
  source      1..3660
              /organism="Lactobacillus delbrueckii"
              /db_xref="taxon:1584"
              452..458
              /note="ribosomal binding site"
              468..3491
              /note="beta-galactosidase"
              /codon_start=1
              /transl_table=11
              /protein_id="AA25240.1"
              /db_xref="GI:149547"

```

Query Match	Best Local Similarity	98.8%	Score 3535	DB 3	Length 3660	
Matches 3590	Conservative	0	Mismatches	0	Indels	45
Gaps	1					
4480	ttggcagccagctctccgagcaatlaatlgaacttggacatggttgcagacaccggtcttgc	4539				
11	ttggcagccagctctccgagcaatlaatlgaacttggacatggttgcagacaccggtcttgc	69				
4540	aagccgaattcgaccacacatggcggtactacggatcgatccgataaaagatttag	4599				
70	-----	-----	-CCGATMAAAAGCTTAGG	85		
4600	cgagcgcttgcgcctggtggtccagacagcgtgaagttctacggcgcatcttgcgcgtactgt	4659				
86	CGACGGCTTTGGCCCTGGTGCACGACAGCGTAAGGCTTAAGCGCAATTTGCCGGTAATCTGT	145				
4660	ccgcagctctgacagacaccggtacatcgatcgctctggaataa tgaacatgsggtcttggat	4719				
146	CGGCCAGCTGGCGCAAGACCCGGCATTCAATCGTCTCGGAATGAACATGGGGGTCTTGGT	205				
4720	cttgatlaaccttggccttgggacacggtlcaaatlaaagggacttggcttgcagctatgt	4779				
206	CTTGATTAACCTTTGGCCTGGGACGATCAATTAACGGGACTTGGCTTGTCAAGTATGT	265				
4780	tgaagagggcagcagcaggtagaagccggccagcagaatctctggaaattcttggagccgggat	4839				
266	TGAAGAGGGCACCCAGGTAGAAGCCGGCCAGAGATCTCGGAATTTGGGACCCGGCGAT	325				
4840	caagcagcccaagccttggagacacaggttaactgttgaacctgcatacaacaggaacttcaac	4899				
326	CAGCAGACGCCAAGCTGGAGACACACGGTAATCTGTGACCGTCAATCAACAGGAACCTTTCAC	385				
4900	aaatagcagatgtctcttgcgcgtacgtgcacacggtlccaagcccttggatgtatcca	4959				
386	AAATAGCCAGATGCTTGTGCGATCGGACACAGCGTCAAGCCCTGGATGATGATTTCAA	445				
4960	gttagaagggaaatattagaataatgaacataatgatttagtaaaaaaagaaggttgacca	5019				
446	GTTTAAAGGGAAGATTGTGMAAATATAGCATTAATTAAGTAATAAAGAAAGAGTTGACCA	505				
5020	ggcagacccgtgccttgcctacgtacacccggaatttacgaagtcataataatcccccga	5079				
506	GGCAGACCTGGGCTTGCTGACTGACCCGGAATTTACGAGATCAATACATTCCCCCGCA	565				
5080	ctccgacacatgactcttccaagacaggaagaacttggagggagggcaagtlccagtttagt	5139				
566	CTCCGACCATGATGCTCTTCCAAAGCAGGAAGAACTCGAGAGGGGGAAGTCCAGTTTACT	625				
5140	gcagtcctccgagcgggagctggtgattgactacgctgaataaacggccagggacacgtcaa	5199				

|||||
Db 626 GCAGCCTCGACGGGGGCGACTGGCTGATGACTACCTGAAAAACGGCGACGAGCACAGTCAA 685
Oy 5200 ctctatgacagaagaagcttgacgatagcaatttgaagtcagatcaagracccgggaacct 5259
Db 686 CTCTATGCAAGAGCTTTGAGAGATGCAATTTTAAGTCAGCAAAAGTAACCGGCAACCT 745
Oy 5260 ggaactgcaagagcttggccagcccaatctatctcaagctccaataatcatatggagcgag 5319
Db 746 GGAAGTCGAAGGCTTTGGCCAGCCCAATATGTCACAGTCCCAATATCCATGAGGAGCGAG 805
Oy 5320 tgaagagatttccgcgcccaaatccaagaanaaalcgcgtcttcttctatgttcagata 5379
Db 806 TGAGAGAGATTTCCCGCCCCCAATTCCAAGCAAAAATCCGGCTGCTTTATGTCAGATA 865
Oy 5380 cttagacctgataagaagcttctgggacaagaagatcagcttgaaagtttcagcgagcggc 5439
Db 866 CTTTGACCTGGATGAAGCTTTCTGGGACAAAGTAAGCTTGAAGTTTGAAGCGGGCGGC 925
Oy 5440 aacagccatctatgtctgagcgagcgccacttcgtgcgtacgaggggagagactcttacc 5499
Db 926 AACAGCCTCTATGCTCTGCTGAAAGGCCACTTGTGCGCTACGGGGAAGACTCCTTTAC 985
Oy 5500 cccaagcgagtttatggttaccgaagttccctcaagaanaaataaacgccttgcagtgagc 5559
Db 986 CCCAAGCGAGTTTATGTTTACCAAGTTCTCAAGAAAGAAATTAACCGCCCTGGCAGTGGC 1045
Oy 5560 tctctacaagatcttctccgcctctgcgtgcgtggaagaccagagactctgcgcagtctcg 5619
Db 1046 TCTCTACAAAGTATCTTCCGCGCTCTGTGCTGGAAGACCAAGACTTGGCCATGTCGG 1105
Oy 5620 ttgttcaagatcagtgactcttccagcccaagccgcgtctgcacttggagagacttaact 5679
Db 1106 TTTGTTCAGATCAGTGAATCTTTCAGGCCAACCCGGCTGTGCACTTGGAGCACTTAACCT 1165
Oy 5680 tacgagcagcttgaacgaataactcaaaaagaagaagcttgaagctcgaagccaatattgc 5739
Db 1166 TACGGCCAGCTTGACCGATACACTACCAAAAAGAAAGTGAAGTGAAGCAAGCANTATTGC 1225
Oy 5740 ctacgcgtctccaaaatgcagactttaaagcttgaagtcggaagtagtgaaggtagctgtg 5799
Db 1226 CTACCGGTTGGCAAAATGCACACTTTAAGCTGGAAGTGGGAGTATGTAAGCTTGGT 1285
Oy 5800 tgcgtcaaaagcttggcccaatcagaagcgagcagctgaacttcaactcgtggttattgccc 5859
Db 1286 TGCTGAAAAAGCTGGGCCCAATCAGAAAGGACAGCTGGAATTCACCTGGCTGATTGCCC 1345
Oy 5860 agtagctgcgtgagcgcggaanaagccttaaccttaacaggtccgcctgtatattatacca 5919
Db 1346 AGTAGCTGCGTGGAGCGCGGAAAAAGCCTTAACCTTACAGGTCCGGCTGTATTATACCA 1405
Oy 5920 ggcagcgagcctctagaagtttagcggaaggaagtggttccgcgaacttgaactaaa 5979
Db 1406 GCGCGGAGACCTTTAAGATGTTACCGCGGAGGAAGTGGGTTCCGCAACTTTGAACATAA 1465
Oy 5980 agacgggattatgaccttaacggccagcgagatgcttcttaaggggcccaacggcgccga 6039
Db 1466 AGAGCGGATTTATGTAATCTTAAAGCGGCCAGCGATGCTCTTCAAGGGGGCAACCGGACGA 1535
Oy 6040 atttgaagtaagttggtgcgggatacaggaagagagatagtatctgggaacatacaagac 6099
Db 1526 ATTGGAAGATTAAGTTGGGTGGGCTATACGGAAGAGATGTATGTATGGGACATCAAGAC 1585
Oy 6100 catbaaagcgaagacaacataatgtctgcgtctgtctcaactaccggaacaaagtcctctc 6159
Db 1586 CATBAACGAAGCAACATCATATGCTGTCCGCTGCTCTACATACCAGAACAGTCCCTCTT 1645
Oy 6160 ttacgggctctgtgacaagtagcgcccttcaagctatgataagcttaaccgggaagcca 6219
Db 1646 TTACGGGCTCTGTACAAAGTAGCGCCTTTACGTATGATGAAGCTTAACCTGGGAAGCCA 1705
Oy 6220 cgggacctgggaaaaaagttggggggcgacgaagatcctaagcttcaatgttccagcgagta 6279
|||||

Db 1706 CGGCACCTGGGAAAAAGTGGGGGGGCGACGAAGATCTAGCTTCAATGTTCCAGCGCATGA 1765
Oy 6280 ccaagatggtctgggaagccaacttaccgggttgaagaacatgagtctgggaaagaa 6339
Db 1766 CCAGCTATGGCTGGGAGCGACGCTTATCCGGGTGAAGAACATGATGGCTGGGACAAAGA 1825
Oy 6340 ccatctcaatcccaactgcttcttaggaatgaagcttcaagcgagctgtcttgc 6399
Db 1826 CCATCTTCAATCCCAATCTGCTTTTAGGCAATAGCTTTACGCCGGCAGCTGCTTTTGC 1885
Oy 6400 ccaaatgctgattacgctcgggaagctgataccgacccgggttccagcataagaagggt 6459
Db 1886 CCAATGGCTGATTAACGTCGGGAAGCGTGAATCCAGCCCGGGTTCAAGCATATGAAGGGGT 1945
Oy 6460 gaaccacaacgggaagtttagagacgcccaagatltgaagcggagatgtatgtccgc 6519
Db 1946 GAACCACAACGGGAAGTTTGAAGAGCCACCGAGTAATTAAGCCGAGTATGCTCCGCG 2005
Oy 6520 caaggtatitgaagaatattgacaaataaaccgaagccatttatctcaagttgaata 6579
Db 2006 CAAAGTTATGAAGATCTTGAACCAATTAACACCAAGCCATTATCTCAAGTTGAATA 2065
Oy 6580 cgtcacgcacatggycaactcgtcgtgagcttgcgcgtacacgycgtcggaaata 6639
Db 2066 CGCTACGCGCATGGGCAACTCGCTGAGTACCTGGCCGCCCTACACGGCCCTGGAAAAATA 2125
Oy 6640 cccccaactacaagggcggtctcaatctggnactgattgaaccaagagactgaaaaagcgg 6699
Db 2126 CCCCACATCAACGAGCGGCTTCAATCTGGAGCTGATTAACCAAGCAAGCAAAAAAGACGG 2185
Oy 6700 gcaactgtcttatggggcgagcttcgatgacggcccaacgacatgaattcttcggaa 6759
Db 2186 GCACCTGCTTTATGGGGCGACTTGCATGATGACCGGCAACCACTATGATTTGCGGGAA 2245
Oy 6760 cgcgtgtcttcttgacccgagactgaatcgcggaacttgcataagtcgaagccctta 6819
Db 2246 CGGCTGTGCTTGTGACCGGACTGATCGCCGAATCGGCTATGTCAAGGCCCTTTA 2305
Oy 6820 cgcacaacttaagtaagaagtaaaagatgggagcctcttccaaaaaagcaaatatt 6879
Db 2306 CGCCAACTTAAAGTAAAGTAAGAAAGATGGGAGCTGCTTCTCAAAAACGCAATTTAT 2365
Oy 6880 taccacaagctcatcttaactcttcttgaagcttcttgaagcttcttgaagcttgaagct 6939
Db 2366 TACCAAGCTCATCTTACTTCTTGTGATGCTTTTGGTCAATGGCAAGTGAACCTA 2425
Oy 6940 ccaagacggcctctgacacttgccttgcgtgcgtgcgaatccggagacttgcctgccc 6999
Db 2426 CCAGAGCGCGGCTCTGAACCTTTGGGCTGGAACCTGGGCAATCCGGACCTTTGGCCCTGCC 2485
Oy 7000 ttggccggaagtcgtgataaagaaggaggtcgtctacccgggtaacggcccaactaaa 7059
Db 2486 TTGGCCGGAAGTCCGTGATGAAAAAGGGGAGGTGCTACCGGGTAAACGGCCCACTTAAA 2545
Oy 7060 aagaagacttgcgttggcgagtgagggtcttcaagcttgaagctgaagcagaagaagtagctca 7119
Db 2546 AGAAGACTTGTGCTTGGGGGAGTGAAGGCTTACGTGTGCTCAACAGAAAGATGACTCA 2605
Oy 7120 aaaactgcgggaatttgaagccggaagggcgagacttgaattgaattcaccactcaact 7179
Db 2606 AAAAGCTGGCGGAATTTAAGCGGGAAGGGCGGCGAGATTATGATTCGCACTCAACT 2665
Oy 7180 aggcctgaagaagaataacttccaaatctcttctccaaggtcaagggtcggcggttcc 7239
Db 2666 AGGCCTGAAGAAATTAACCTTCCAAATTCCTTCTCAAAAGGTCAAGGGCTGGCGGGTTTC 2725
Oy 7240 cctcaagatagcgttagggaataacttgaagcggtcgcgggaatttacccttcggcgggc 7299
Db 2726 CCTCAAGTATGCCGGTGAAGGAATCTTGAAGCGGCTGCCGAATTTACCTTCTTGGCGGCG 2785
Oy 7300 cctgaacgacaacgacccgggagcgtgttagcgatgatcttgcgcgggtgggaaatgac 7359
Db 2786 CCTGACGAGCAACGACCGGGAGCTGGTTACGGCTATGATCTGCGCGGTGGAAAAATGCG 2845
|||||

QY 7360 cggcaagatagcccgcttgaaagacatcagctgcgaaggtcaaggaagaagctccgttttgt 7419
|||||
Db 2846 CGGCAAGATATCCCGCTTGAAAGACATCAGCTGCAGAGTCAAGGAAGACTCCGTTTGGT 2905
QY 7420 caaagatcccttcgctgcgtgcgttccttaaggtatgatttaaccgttgactatgaat 7479
|||||
Db 2906 CAAGACTCCCTTTACGTTGCGCTGCTTAAAGGGGTATTTAACCGTGACCTATGAAGT 2965
QY 7480 cgaatgagcgggcaagatctgtatcaacagcttaccacagcgcggaagaagctgtct 7539
|||||
Db 2966 CGATGAGAGGGGCAAGATTGCTGTAAACGTACTTCCAGGCGGCAAGAAAGCTGCT 3025
QY 7540 ctgccaagcctttgcttgtaacatcgtgcccctgcgaagaagctgacgaattaccgttacta 7599
|||||
Db 3026 CTTCGACACTTTGGCTTGAAACCTGCGCCCTGCCAAGAAGTACGATTACCGCTACTA 3085
QY 7600 tgcgttcggagcctaataatgagcttaccacagcgcgttggaaggttaattaccctggcatct 7659
|||||
Db 3086 TGCCTGGAGCTTAATGAGAGCTACCCAGACCGCTTGGAAGGTAAATTAACCTGGCATCTA 3145
QY 7660 ccaaggaacggttaaaaaagaaccttagcccatatcgtccgcgagaaacgggaaccgag 7719
|||||
Db 3146 CCAGGAGACGGTAAAAAGAACTTAAAGCCATATGCTCCGCGAGAAACGGGCAACCGGAG 3205
QY 7720 caaggttcgttgtaacagcttcttgatgaaagggcggttggaatttacgccaatg 7779
|||||
Db 3206 CAAGTTCGCTGCTACAGCTCTTGATGAAAGGGCGGCTTGGAATTACGCGCAATG 3265
QY 7780 ggcagacttgtaactgtctgtcttccatattctgcgcgcccgaattgaaagcagcgagca 7839
|||||
Db 7326 GCGACACTTGAACCTTGCTGCTTCCATATCTCCGCCGCCCAATTTGAAGCGCGACCA 3325
QY 7840 cgctttagactgaactaacaattacacttgggttagaagccttaagcgcccgaatggaggt 7899
|||||
Db 3326 CGCTTTAACTGACTTAACAATTACACTTGGGTAGAGCCTTAAAGCCGCAATGAGG 3385
QY 7900 cggcgaggatgactctctgggggcagaaaggtccaccgcgaattctcgtgagctcaaaa 7959
|||||
Db 3386 CGGCGGAGTATGATCTCGGGGCAAGAGGTCCACCGGAATCTCGCTGATGCTCAAAA 3445
QY 7960 agccgcagactgcgtgcgtgattcagcccttactaataaataatgataatgac 8019
|||||
Db 3446 AGCCGCGAGCTCCGCTGCTGATTCAGCCCTTTACTAAATATGCTCAATTTGAC 3505
QY 8020 ttaacagatgaattttagtaaaagcaagcgagtggaagatggaacatcaaga 8079
|||||
Db 3506 TTACAGATGAATTTAGTAAAGCAAGCGAGTGAAGAGATGCAACGATCAGAGA 3565
QY 8080 agtgcaagcagcgcgctgtcgtacgagcgt 8114
Db 3566 AGTGCAAGGCAAGCGCGGCTGCTGCTAGGACG 3600

RESULT 2
AX009488 5059 bp DNA PAT 06-SEP-2000
LOCUS Sequence 1 from Patent WO961627.
DEFINITION AX009488
ACCESSION AX009488
VERSION AX009488.1 GI:9996771
KEYWORDS
SOURCE
ORGANISM
Lactobacillus delbrueckii subsp. bulgaricus.
Lactobacillus delbrueckii subsp. bulgaricus.
Bacteria: Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
REFERENCE
1 (bases 1 to 5059)
AUTHORS Brignon,P., Gendreau,F. and Benbadis,L.
TITLE Mutant lactobacillus bulgaricus strains free from beta-galactoside activity
JOURNAL Patent: WO 961627-A 1 02-DEC-1999;
GENEVAIS DANONE CO (FR); BRIGNON PIERRE (FR); GENDRE FRANCOIS (FR); BENBADIS LAURENT (FR)
FEATURES
Location/Qualifiers

source 1..5059
/organism="Lactobacillus delbrueckii subsp. bulgaricus"
/db_xref="taxon:1585"
CDS 122..1873
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07487.1"
/db_xref="GI:9996772"
/translation="MIFITMLITAIRIGEVLLDPLIGNADITRESRWKFKPWWVG
GISLALAFDTEFGINOSNPVYLVIFGIYVLMIDIEFSKDTGFWAMIPLSLD
SREERKSTFARVSTIGANLVGVITPILIFSASRANPNRGGDQNGFEFLIAYV
ILSTITVGLGHEVKSARSENKETTLLKQVYCONDOLMLIAPAVYGLINTLN
ALQVYFSYILGDARGSLTYINTFVGLISASFPSPDLAKFNRRRLPYACTAYMLG
IGPVSASGSLALVGAFFFIQPLAFIVLMIISDAVFGQLKTRDEALITLV
RPLVDKLGALSNMFVLIALTAGMTGACSTTAHGOVFKLAMEALPVMILIAV
SIFAKVFLTEEKAEIYDQETOFESGHAOKPAQAEFTLASPVSGLMIDWDVDP
VFADKXIGDFALPADGKVVYAPFAGTYROIARLTHRSITVLENEGLVLIHLGGTVK
LNGTFPVSVEEGSOVEAGQOILEFWPAIROALTDIVITYITVINSEFANSOHLPT
GHSVQALDVERKLEKN"
1877..4519
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07488.1"
/db_xref="GI:9996773"
/translation="MSNKLVEKRVQADLAMLTDPEYEVNTIPRSHDSFOSOE
LEEKSLVQSLDDMLIDYVLENGQVPVFAEDFDSDNFESVPGMLLEQGGOP
VYVQYDPDSEIEFPQIPSKNPLASVYFEDPEADMEKESLKEFGATATAYVWL
NGHVFVGDEDSFPTSEPMVTKFLKKNRNLVALYKYSASMLDODMEKGLFRSY
TLQAKPRHLDELKLTASLTNYOKKLEVAENIATRIPLNASFLEVDSGDIYAEK
LGPIRSEQLFTLADLPVAWSAKPNLYOVLVLYQGSLLFVRSQVGRNRELK
GYIMVNGRIYVEKANSRHPDSKLGRLATTEEDMDIKTKMRNINAVRCSHYNOL
FYRLCDKXGLYVIDEANLHSGTWKVGHEDPSENVGDDQHMIGASISVKMMAR
DKNASTILISLGNESYAGTYEAMADYVRADPTRVQHYGVTHNRKEDATQIESR
MYAPAKVIEETLTKNPKAPRTISVERAHMGNSVDLIAATLFEYYPHROGGFTIMID
QGLEKDHLLTGGDPDPTDYERICNGLYVADTRESKLANVAILANLLEKVDGQ
LELNNDLLETSSSYFLTSLVYDGLVQSRPLFGLEPESGTFALPVEVADENK
EVVAVRVAHLKEDLPMADEGPTVAEVAQKLPEFKGEPDSDVSYNLGLNNF
OILFSKVGMPVSLKTAGREYLKRLPEFTFWRALTDNRGAGYGDILARWNAKAYAR
LKDLSKYEKEDSVLKTAFPLPVALKGLVITYEVDGKGLAVTADFGABEAGLLPA
FGMLALPKELTDVRYIGLGNES"
BASE COUNT 1229 a 1341 c 1323 g 1166 t
ORIGIN

Query Match 43.1%; Score 3497.2; DB 9; Length 5059;
Best Local Similarity 99.5%; Pred. No. 0; Mismatches 13; Indels 3; Gaps 1;
Matches 3518; Conservative

QY 4584 ccgataaaagttaagcgagcagcgttgcctgtgtgcacagcagcgttaaggtctacgcgc 4643
Db 1479 CCGACAAAAGTTAGCGCAGCGGCTTGCCCTGCTGCGACACACGCTAAGTGCTACGGCC 1538
QY 4644 catttgcgtaactgtccgcagcttgcgaagaaccgcgacatcgatcgtctgtaaatgt 4703
Db 1539 CATTTCGCGTACTGCTCCGCCAGCTGGCCAAACCGCGACCTGATCTCGTGGAAATG 1598
QY 4704 aacatggggtctgtgtctgtatctgaatccttgcccttgcccttgcccttgcccttgccct 4763
Db 1599 AACATGGGGTCTTGCTGTGATTCACCTTGCGCTGGCGCGGTCAATTAATTAACGGGACTG 1658
QY 4764 gcttgcagcatttgaagaggcagcgaagtagaagcggcgccagcagcagcagcagcagc 4823
Db 1659 GCTTTGTCAGCTATGTGTAAGAGGCGACGAGTGAAGACCGGCGCAGATCTCGAAT 1718
QY 4824 tctgggaccggcgatcaagcagcgaagctgagcagcagcagcagcagcagcagcagcagc 4883
Db 1719 TCTGGGACCGCGGATCAAGCAGCGCAACGCTGAGACACGCTAATTCGACCGCATCA 1778
QY 4884 acaaggaacattcaacaataagcagatgcttcttcgcagtcggcacaagcgtccaagcc 4943
Db 1779 ACAGGAAACTTTCCCAATTAACCGAGATGCTCTTGCCGATCGCCACACGCTCAAGCCC 1838

OY	4944	tgratgatyatltcaagcttgaaggggaagaaatagaaaatgacataagtlagtaaaag	5003
Db	1839	TGGATGATGATTCACAACTTGAAGGGAGAAATTAGAAAATGAGCAATAGTTAGTAAAG	1898
OY	5004	aaaaaagattgaccagagagacttgcttgcttgacttgaccgtgaccggaatttagaagfca	5063
Db	1899	AAAAAGATTGACCCAGGACGACTTGACCCTGGCTGACTACCCGGAACTTACGAAAGCA	1998
OY	5064	atacaatltccccgcacgtccgcacatgagltccttccaaagccaggaagaacttggagag	5123
Db	1959	ATPACAATTCCCGCAGCTCCGACCATGAGTCCTCCAAAGCCAGGAAGACTGGAGAGG	2018
OY	5124	gcaagtcagttttagtgacgttcccttgacgggagactgagctgtatgtactaacgtgaag	5183
Db	2019	GCAAGTCCAGTTTAGTGCACATCCCTTGAGCGGGACAGCGTGATTACACCTGAAACG	2078
OY	5184	gccaagagccaagtcacacttctatgcaagaagacttgcagatgacatttgaatgacga	5243
Db	2079	GCCAGGAGCCAGTCACTTCTTATGCAAGAACTTTCACATAGCAATTTTAAGTCAGCA	2138
OY	5244	aagtaaccggaacacttggaaactgcaagacttggccagcccaagtatgtcaacgtlceat	5303
Db	2139	AAGTACCCCGGCACTGTGAAGCTCAAGGCTTTGGCAGCCCAATGTCAACGTCACAT	2198
OY	5304	atccatggagagcgagatgagagagatlttccggcccacaaatcccaagcaaaatcgctcg	5363
Db	2199	ATCCATGGAGCGGAGTGAAGAAATTTTCCCGCCCAATTTCCAAAGCAAAAATCCGCTCG	2258
OY	5364	ctctctatgacagataacttggaccttggatgaagcttctcggagacaagaagtcaagctga	5423
Db	2259	CTTCTTATGTCAGATTACTTGTGACCTGGATGAGAGCTTTCGGGACAGGAGTCAGCTTGA	2318
OY	5424	agtttagccggggggcgcaagccatctatgcttgcttgaagggcaacttgctgctacg	5483
Db	2319	AGTTTGAAGGGGGGGGCAACAGCCATTAATGCTGGCTGACAGCGCACTTCCTGGCAGC	2378
OY	5484	gggaagacaccttaccaccaagcgagtttatgttaccagacttctcgaagaagaagaata	5543
Db	2379	GGGAAGACTCCTTATCCCCCAAGGAACTTATGCTTACCAAGTCTCTCAAGAAACAAATA	2438
OY	5544	acgcgcctggcagtgagctctctacaaglatcttcgcgcctcctgctgtgaagaccaagact	5603
Db	2439	ACCCCTCGGAGAGTGCGCTCTACACAAATTTCTTCGCGCTCGCTGAGAACCAAGGACT	2498
OY	5604	ctctggccatgctcgtgtgttctcaagtaagtgaaacttcaaggccaaagccgcgcttgact	5663
Db	2499	CTCTGGCCATGCTCTGGTTTTCAGATCAGTGACTCTTAAAGGCCAAGCGCGTCTGCACT	2558
OY	5664	tggagagaccttaagcttcaagccagacttgcacgataacttccaaagaagaagcttggag	5723
Db	2559	TGGAGGACCTTAAAGCTTACGGCAGCTTGACCGATTAACACAAAAGAAAGCTGGAGG	2618
OY	5724	tcgaagccaatattgctactacgcttgcacaatgccaagctttagcttgaagtgcgggata	5783
Db	2619	TCGAAGCCAAATTTGGCTTACCGTTCACAAATGCCAGCTTTTAAGCTGGAAATGGGGATA	2678
OY	5784	gtgaaggtgaacttgatgtctgcgaaagcttgggcccacatcgaaagcagcagcttgaatca	5843
Db	2679	GTTGAAGGTGACTTGTTGCTGTAAGAAAGCTGGGCCAATCAAGAACGACGACTGGAAATCA	2738
OY	5844	ctctgtgctgatttgcagtagcttgctgtgaagcgcggaagaagcttaacttaccaggtcc	5903
Db	2739	CTCTGCGGTGATTTGGCCAGTGAAGCTCTGGAAGCGGAAAAAGCTTAACCTTTACAGGTCC	2798
OY	5904	gctcgatattatcccaagcagcagcctcttagaagttagccggagaaagtgggtttcc	5963
Db	2799	GCTCGATTTTATACACAGCGAGGAGCCTTTTAGAGGTTTACCGCGAGAAAGTGGGTTTCC	2858
OY	5964	gcaactttagactaaagaagcgagatatgtacttcaagcgccagcgagatcgltcttcaag	6023
Db	2859	GCAACTTTGAACATAAAGAGCGGATATATACTTTAAGCGCCAGGAGATGCTTTCAGG	2918

QY	6024	gggccaaecggcgcgaagtttggacaagtcgaagtttgggtgtatctacggaagaagata	6083
Db	2919	ggggccAACGGGCACGAATTTTGACAGTAAGTTGGGGCGGCTATCCAGAGAGGATTATGA	2978
QY	6084	tcctggacatcaagaacataaagaacgaacaaataaagtcgtccgcgtcctcaacc	6143
Db	2979	TCCTGGGATATCAACATCATACATACAGCAACATCATCTCTCCGCTGCTCTCAAC	3038
QY	6144	cgaaacagtcctccctcttttacggcctctgtbaacaagtaagccttaagtcacttga	6203
Db	3039	CGAACAAGTCCCTCTTTTACCGGCTCTGTACAGTAAGTACGGCTTTACGTCAFTGATGAG	3098
QY	6204	ctaacttggaaaagccacgycacctgggaaaaagttggggggcagcaagaatccctagctca	6263
Db	3099	CTAACCTGGAAAGCCACGCACTTGGGAAAAAGTGGGGGGCAGCAAGATCCTAGCTTCA	3158
QY	6264	atgtttcaaggagatgaccagcatctgtgtggggccagctctaccgggtgaagaacatga	6323
Db	3159	ATGTTCCAGGGGATGACACACATTTGGCTGGAGCCAGCTTATCCGGGTGAAGAACATGA	3218
QY	6324	tgagctcggaagaacaaacagcttcaatccaaatcagtcgtcttaggaatagtcctacg	6383
Db	3219	TGGCTCGGGACAGAACCATGCTTCATATCTGATCTGGTCTTTAAGCATAGTCTTACG	3278
QY	6384	ccggacactgtctcttggcccaaatgctgtatcagctccggaagcgtatccgaaccgggttc	6443
Db	3279	CCGGACGTCGTTTGCCCAATGGCTGATTAGCTCCGGAAGGCTGATCCGACCCGGGTTC	3338
QY	6444	agcacatgatgaagggtgtaccccaacaaccggaagtltgaagcagccaccagattgaagcc	6503
Db	3339	AGCACTATGAAGGGGTGACCCACACACGGGAAGTTTGACACACCACCCAGATTGAAACC	3398
QY	6504	ggatgtatgtcccggaagtgatattgaagaatccttgaacaaataaaccaacgaagccat	6563
Db	3399	GGATGTATGCTCCGGCCCAAGGTAATTGAAGATACTTGCATTAACCATACCAAGCCAT	3458
QY	6564	ttactcagttgaatagcctcagccacatgggcacatccgtctggttaccctggccgctaca	6623
Db	3459	TTATCTCAGTTGATACGCTACAGCCATGGGCACTCGTGGTACCTGGGCGCGCTACA	3518
QY	6624	cggcccttggaaaatatcccccaactaccagggcggtctcaatcttggactgtatgaccaag	6683
Db	3519	CGGCCCTGGAAAAATATCCCCACTACACAGGGCGGCTTCACTCGGGACAGGATTGACCAAG	3578
QY	6684	gacgtgaaaaaagacgggcacctgtcttatatggggggagacttgaataccggcaacagact	6743
Db	3579	GACTGGAAAAAGACGGGCACCTCTTATATGGGGGCACCTTCATATACGGGCCAACCGACT	3638
QY	6744	atgaattctgcgggaaagcgcgtgtctcttgcctbaacggagctaaatccggcaaacctgcta	6803
Db	3639	ATGAATTCTGGGGGAACGGCGCTGTCTTGTCTGACCGGACTGAATCGGCCAAACTGGGTA	3698
QY	6804	atgtcaagggcccttiacgccaaccttaagttagaagtaaaagaatggcagctcttccctca	6863
Db	3699	ATGTCAAGGGCCCTTTAGCCCAACTTAAGTTGAAGATAAAGATGGCAGCTTCTTCYCA	3758
QY	6864	aaaaagcaatttatattaccaagaagctcaatctactactcttgcactgtcttggctgcg	6923
Db	3759	AAAAAGCAATTTATTTTAAACCAAGCTCATCTTACTCTTGTGACTGTCTTTTGGTCG	3818
QY	6924	atggcaagttgaacttacaagaagccggcctctgaaccttggccttgaagccttggcgaatccg	6983
Db	3819	ATGGCAAGTTGACCTTACACAGCGCGGCTCTGACCTTTCGCTGGAAGCTTGCCAAATCCG	3878
QY	6984	ggaccttggcccttgccttggcgggaatgcgtgtatgtaaaaggggaagtcgtctacggg	7043
Db	3879	GGACCTTTGCCCTTGCCCTTGCGCCGAAGTGCATGATGAATAAAGAGAGGTCTCTTACCGGG	3938
QY	7044	taagcgcccaactaaagaagaagcttgccttggcggaatgaagggcttcaactgttgcctgaag	7103
Db	3939	TAAAGGCCCACTTTAAAGAAAGACTTTCCTTTGGGGGATGAGGGCTTCACTGTGGCTGAAG	3998
QY	7104	cagaagaagatgactcctcaaaagcttcggaattttaaagccggaagggcgccgaatttattg	7163

|||||
Db 3999 CAGAGAAGTGTAGCTCAAAAAGCTGCGGAATTTAAACCGGAAAGGGGCGAGATTGATTTG 4058
Oy 7164 attcgaatacaacctgaagcctgaagaataaactccaattctctcccaagtga 7223
Db 4059 ATTGCGACTACACCTAGCCTGGAAGGAATTAATCTCCAAATTTCTTCTCAAGGTCA 4118
Oy 7224 agggctgagcggcttccctcaagatgacggtaaggaaactgaagcggctgcggaat 7283
Db 4119 AGGGCTGGCCGGTTTCCCTCAAGTATGCCGTAGAGGAATACTTGAAGCGGGCTGCCGAAT 4178
Oy 7284 ttacctctgagcggcctctgacgaacagaccgggagcgtggttaagcgtatgctcgg 7343
Db 4179 TTACCTTCTGGCGGGCCCTTGACGGAACAGACCGGGGAGCTGCTTACCGCTATGATCTGG 4238
Oy 7344 ccgggtggaataatgcggaagatgcccgcctgaagaacatcagccgagagtcgaag 7403
Db 4239 CCCGTTGGGAATAATGCCGGAAGTATGCCCTTAAAGACATCAGCTGCGAGGTCAAG 4298
Oy 7404 aagactcggcttggtaagaactgaccttaacgtgtgcctgtcgaagggtgattaa 7463
Db 4299 AAGACTCGGTTTGGTCAAGACTGCTTACGTTGCTGCTGCTTAAAGGGTGATTTAA 4358
Oy 7464 ccgtgacctatgaagtgcgtgacgggcaagatgtcgttaacagctgactccagcg 7523
Db 4359 CTGTGACCTATGAAGTGCATGAGCGGGCAAGATTGCTTAACAGCTGACTCCAGCG 4418
Oy 7524 cgggaagaagctggtctctctgcaagccttgctgaacctgcccctgcgaagaagatga 7583
Db 4419 CGGAAGAAGCGGGTCTTCTTGCGAGCTTTGGCTTGAACCTTGAGCCCTGCGCAAAAGACTGA 4478
Oy 7584 ccgaataccgactactatgctctggaacctaaatgaagatgaccagaccgcttggaagta 7643
Db 4479 CCGATTACCGGTACTAGTGTCTGGGACCTTAATGAGAGCTAACCAAGCCGCTTGAAGGA 4538
Oy 7644 attactggtgcatctacgaaggagcggttaaaaaaacttaagccaca---tcgtccgc 7700
Db 4539 ATTACTGTGGGATCTAACGAGGAGCGGTAAAMAAACCTTAAGCCATTAACGTCCGCGC 4598
Oy 7701 aggaagaagggaacccgggcaaggtctgctgttaccagctctcttgataaaggcggt 7760
Db 4599 AGGAAGCGGCAACCGGAGCAAGGTTCCCTGTGTACCACTCTTTTATGAAAAGGCGGCT 4658
Oy 7761 tgaatactacgccaatgagggaagacttgaaactgtctgcttgccatattctgcgcgc 7820
Db 4659 TGAATTTACGGCCCAATGCGGAGACTTGAACCTGTCTGCTTGCATATCTGCGGCC 4718
Oy 7821 aaattgaagcagcggaaccaagcctttgaactgactaacaattgaacttgggttaagcct 7880
Db 4719 AAATTGAAGCAGCGGACACGCTTTTGAACGTATACATAATTAACCTTGGGTTAGAGCTT 4778
Oy 7881 taagcgcccaatgggggtcgcgcgaggatgaactccttgaggggcagaaggtccaccggaat 7940
Db 4779 TAAAGCGCCAATGGGGGTCCGCGGGGATGACTCCTGGGGGAGAGAGTCCACCGGGAAT 4838
Oy 7941 tctgcctgtagtctcaaaaagcccgccagccttgcgctgtgattcaagcccttactaa 8000
Db 4839 TCTGCGCTGATGCTCAAAAAGCCCGCAGCTCCGCTGCTGATTCAGCCCTTTTACTAA 4898
Oy 8001 aataaatgtctcaatgtacttaacaggaatgaattttagtaaaagcaagcgagtgaagga 8060
Db 4899 AATAAATGCTCAATTTGACTTAACAGGATGAATTTTGTATAAACCAAGCAGTGAAGA 4958
Oy 8061 agatggcaagatcagaagaatgcaaaagcagcgcggtgtcgtcagcagcggt 8114
Db 4959 AGATGGCAGACATGACAGAACTGCCAAGGACGCGGCTGTCTGCTAGCAGCGGT 5012

RESULT 3
LBALACZBUL 5015 bp DNA BCT 26-Apr-1993
LOCUS L.delbrueckii bulgaricus lactose permease and beta-galactosidase
DEFINITION (lacZ) genes, complete cds.

ACCESSION M55068 M38754
VERSION M55068.1 GI:149564
KEYWORDS beta-galactosidase; lactose permease.
SOURCE L.delbrueckii bulgaricus DNA.
ORGANISM Lactobacillus delbrueckii
Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
REFERENCE 1 (bases 1 to 5015)
AUTHORS Leong-Morgensthaler,P.M., Zwaalen,M.C. and Hottinger,H.
TITLE Lactose metabolism in Lactobacillus bulgaricus: Analysis of the
JOURNAL primary structure and expression of the genes involved
MEDLINE J. Bacteriol. 173, 1951-1957 (1991)
FEATURES
source
1..5015
Location/Qualifiers
/organism="Lactobacillus delbrueckii"
/sub_species="bulgaricus"
/db_xref="taxon:1584"
/feature_1db="ATCC 11842"
102..1985
/gene="lactose permease"
102..1985
/gene="lactose permease"
/note="putative"
/codon_start=1
/transl_table=1
/product="lactose permease"
/protein_id="AA2524.1"
/db_xref="GI:149565"
/translation="MKKLVSRSLYAAAGANDVYATLSTFYFVETTHLPFANADHK
MFIITNLITAIRIGEVLDLPJLGNALIDRESRWGRKRPVWVGGLISLALFLPD
FGJNOSKRPVYLVIFGIVYLIMDIFFSFKGTGPMWMI PALSDRSREKRTSPARG
STIGANLVGVITPILIFESAKANPMDGKGFFALIVAVIGILSIITGLSTHEV
KSALRESNEKTTLOKRVFGONOLMLAAYWYFGVIGINTLALDLYRSYITLGA
RGVSLITVTFVGLISASPPRSIAKKNRRLRLACIAVWLIGIVSAVSGSLAS
LVGAFFPTPLFLVYLMITSAVERGQKTHRBALITSRPLVDKGLSWM
FVSLIALTAGMTTATSTTTHAOMYKFLAMEFLPVMILIAVSTPAKVFLEKH
AEIVDOLETQFGSHAQKPAQAESEFTLASPVSGOLMIDWDVDFADKIKGDFALV
PADGKVVYAPFAGTIVROLAKTRHSIVLENEHGLVLIHIGLGTAKLNGTGFVSVEGS
OVEAGQOILFEWMDPAIKQAKLDIVIVTVINSETFANSOMLPIGHSVQALDVFKLE
GKN"
1989..5015
/gene="lacZ"
1989..5015
/gene="lacZ"
/note="putative"
/codon_start=1
/transl_table=1
/product="beta-galactosidase"
/protein_id="AA2524.1"
/db_xref="GI:149566"
/translation="MSNKLVEKRVDOADLMLTPPEYEVNTIPPHSDHESFQSOE
LEEGSSIVOSLDGDMLIDVAENGQGVNFAEDFDENFSKVSVPNGNLEQFGGO
YVNYQYPMDSSELEFPQIIPSNKLASYRVYFDIDEPDAPMKVEVSLKPDGATAYVM
NGHVVGEDSEFTPSEFNVTKFLKENRLVALIKTSSASWLEDQDFWRSGLFRSY
TLOAKPRHLEDKLTASLIDNYOKGLEVPANLAIYRLPNASFLEVRDSEGLVAEK
LGPRLSEOLEFTLADLPVAAMSAEKPMLYQVRLTLYOAGSLLEVSROEVRNREL
GIMVLSORAIYFECANRHEPDSKIGRAITEDMDIMDKTMRSMININVCSHYNOSI
FYRLCDXGIVYVIDEANTLESHGWEKXGKSPFNPYGGDOHMLGASLSVKMMAR
DKNNAIILINSIGNESTAGTYFAOMADYVRKADTPROHTEGYTHNKPFDATOBESR
MYAPAKVIEEYLTNKPAPKPTISVEYAHAMGNSVDLAAVTALEKTPHTQGFIMDID
QGLKEDHLLYGGDFDRPDYEPFCNGILYFADTESPKLANVALYANLKEVDKQ
LEKNDMLFTWSSSYEFLTSLVDGKLTYSRPLTEGLEPESGTFALPPEVDEKG
EYVYRVANLHKEDLPMADEGTVVAEEVIAOKLPEFKEGRLDVLSDVNLGLKNNF
OIFESKYGMPVSLKYAGREYLRKLPEFTFRALTDMDRAGVGYDLARWENACKYAR
LKDISCEVKEDSVLYKTAFTLPALKEDLVITYEVDORAKIAYADPFGAEAEALLPA
FGMLALPKELFTDRIYVGLPNESTYPRLEBNYGLTYOGAAKKNFSPYLRQEGNS
KVRWYOLFDEKSGLEFANGADLNLISALPYSAOIEADAHAFELTNNYTWVRALSAOM
GVGDDSDMGQVHPEPCLDOKAROLRLVIOPLLLK"

BASE COUNT 1231 a 1322 c 1292 g 1170 t

ORIGIN

Query Match 41.7%; Score 3386.6; DB 3; Length 5015;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3408; Conservative 0; Mismatches 14; Indels 3; Gaps 1;

QY 4584 ccgataaaaaagttaagcgcagcgtcttgccttggctgcgaagcagcaggtlaagctlacgcgc 4643
|||
DB 1591 CCGACAAAAAGTTAGCGAGCGGCTTTGCCGTGGCAGCAGACGGTAAGCTTACGCCGC 1650
QY 4644 catttcggttactctccgcagcttgcccaagaccgcgcagctcgatcgtctctggaatg 4703
|||
DB 1651 CATTTCCGGTACTGTCGGCAGCTGGCCAAAGCCGGCAGCTGATCTCTTGAAATG 1710
QY 4704 aacatgggtcttggtcttgattcaccttgcttgccgagcagctcaaatlaaagcagctg 4763
|||
DB 1711 AACATGGGCTTGTGCTTGTATTCCTTGGCTGGCAGCGCAATTAAGGGAGCTG 1770
QY 4764 gctttgcaagctatgttgaagagggcagccaggtlaaagccgcgcagcagctctgnaat 4823
|||
DB 1771 GCTTTGTCAGCTATGTTGAAGAGGGCAGCCAGGTAGAGCCGGCCAGCATCTGGAAAT 1830
QY 4824 tctgggaccgcgcgcatcaagcagccagctggaacagcagctgaatcgtgacgctcaatca 4883
|||
DB 1831 TCTGGGACCCGGGATCAAGCAGGCGCAAGCTGGACGACGGTAATCTGATCCGTATCA 1890
QY 4884 acagcgaaacttcacaaatagccagatgctcttgcgcatcgcgcacagctcaagccc 4943
|||
DB 1891 ACAGCCAACTTTGCGAATAGCCAGATGCTCTTGGCGATGGCCACAGGCTCCAAAGCCC 1950
QY 4944 tggatgatgtatcaagcttgaagaggaagatagaataatgaacaaatgaatgaatgaatga 5003
|||
DB 1951 TGGATGATGATTTCAAGTTAGAGGAGAAATTAGAAAAAGACAAATTAATTAATGAAG 2010
QY 5004 aaaaaaggttgcagcagcagcagcctgctgctgctgactgacccggaaagtttacgaatga 5063
|||
DB 2011 AAAAAAGAGTTGACGAGCAGACCTGGCCCTGGCTGACTGACCCGGAGATTAGAGAGTCA 2070
QY 5064 atacaattccccgcgacactccgacactgaagctccttccaaagccaggaagaaacttgaagag 5123
|||
DB 2071 ATACAAATTTCCCGGACTCCGACCATGATGCTTCCAAAGCCAGAGAAAGACTGGAAGAG 2130
QY 5124 gcaagtcagctttagtgcagctcccttgaacgggagcagctgctgactgactacagctgaagag 5183
|||
DB 2131 GCAGGTCAGATTAGTGAAGTCTTGAGCGGGAGCTGGCTGATGACTAGCTAGCGTAAGAACG 2190
QY 5184 gccagggaccagctcaactctctacggaagacttgaacagtagcaatlttaagtcagta 5243
|||
DB 2191 GCCAGGACCACTGCACTTATGCAAGAGACTTGTGACATAGCAATTTTAATGCTAGTCA 2250
QY 5244 aagttaccgggaacctggaactgcaagcgttggcgaagcccgagctatgttaagctcaat 5303
|||
DB 2251 AAGTTACCCGGCAACTGGAAGCTGAAAGCTTTGGCCAGCCCAAGTATGTAACGTCCAAAT 2310
QY 5304 atccatgggagcagctgaaggaatttccgcgcccaaatctcaagcaaaatccgcgtcg 5363
|||
DB 2311 ATCCATGGGAGCGGCACTGAGAGATTTTCCCGGCCCAAAATTCACAGCAAAAAATCCGCTCG 2370
QY 5364 ctctctatgtcagaatacttgcagctggaagagccttctggaacaaagaaatgaacttga 5423
|||
DB 2371 CTCTCTATGTCAATATCTTGTGACCTGGATGAACTTCTTGCGAAGGAAGTACAGCTTGA 2430
QY 5424 agtttgaagggggggaagcaactatgtctgctgctggaagccgcaactgctgtgctcag 5483
|||
DB 2431 AGTTTACGGGGGGGCAACGCACTATGTCTGGCTGGAAGCGCCACTTGTGTGGCTACG 2490
QY 5484 gggaagaactccttaccacaaagcgaattttagtltacaaagcttcccaagaaagaaata 5543
|||
DB 2491 GGGAAAGACTCTTTACCCCAAGGAGTTTATGTTACCAAGTTCTCTCAAGAAAGAAACA 2550
QY 5544 accgcttgagagtgctctctctacaaatattctctcgccctctctgctggaagcagagact 5603
|||
DB 2551 ACCGCTTGAGAGGCTCTCTACAAATATTTCTTCCGCTTCCTGCTGGAGGACAGCAAG 2610
QY 5604 tctgggcatgtctggttctgttcagatcagtgactcttcaagcgcaagccgcgtctgact 5663

DB 2611 TCTGGCCATGTCTGTTTCTTTCAGATCACTGACTCTTCAAGGCCCAACCCGCTGTGACT 2670
|||||
QY 5664 tgaagaccttaagcttacgcgagcacttgaccgataactacaaagaaagaaagcgtggaag 5723
|||
DB 2671 TGGAGAGCACTTAAGCTTACGGCAGCTTGACCGATTAATCTCAAAAAAGAAACCTGGAG 2730
QY 5724 tgaagcaaatattgtcctaccgcttgcgaatgcccagcttlaagcttgaagtgcgagata 5783
|||
DB 2731 TCGAAGCAATATTTGCTTACCGCTTGCCAAATGCCACCTTTAAGCTGGAAGTGGGATA 2790
QY 5784 gtgaaggtgacttggtctgctggaagaaagctgggcccaatcagaagcagcagctgaatca 5843
|||
DB 2791 GTGAAGGTGACTTGTGTTGCTGAAAAGCTGGGCCCAATCABAAGCAGCACTGGAATCA 2850
QY 5844 ctctgctgatttgcgcaagtagctgctctggaacgcggaagaaagccttaacttaccagttcc 5903
|||
DB 2851 CTCTGGCTGATTTGCCAGTAGTGTCTGAGCGCGAAAGCCTAACCTTTACAGGTTCC 2910
QY 5904 gctctgtattataccagggcagcagcctcttaaggttgaagctggcggaagtggtttcc 5963
|||
DB 2911 GCCTGTATTTATACAGGCAAGGCAAGCTCTTAAAGTTACCCGGCAGGAAGTGGGATTC 2970
QY 5964 gcaacttgaactaaagagcaggaattatgtaccttlaacgcgcagcagctgcttcaag 6023
|||
DB 2971 GCAACTTTGAACTTAAAAAGAGGGATTTATGACTTAACGGCCAGCGATGCTTCAAG 3030
QY 6024 gggccaaaccgcaagaaatttgaacagtaagtttgggtcgggtatcaacggaagagataga 6083
|||
DB 3031 GGGCAACCGCGCAGAAATTTGACAGTAAGTTGGGCGCGGTATCACGAAGAGATATGA 3090
QY 6084 tctgggaacatcaagaagcaatgaagcgaagcaaatcaatgtctgtccgtcttcaactacc 6143
|||
DB 3091 TCTGGGATATCAAGACCATTAACCGAAGCAATCAATGTCTGCCGTCTCACATACC 3150
QY 6144 cgaacagctccctcttcttaccgctctgtgacaagtaacgcgccttlaagctatgataag 6203
|||
DB 3151 CGAACCACTCCCTCTTTTACCGGCTCTGTGCAAGTACGGCCTTTAGCTATGATGAAG 3210
QY 6204 ctacacttgaagaaacacagcagcacttgggaagaaagttggggggcagaaagatcctagcttca 6263
|||
DB 3211 CTAACTGGAAGCAACGCGCACTGGGAAAAAGTGGGGGGCGCAAGAAATCTAGCTTCA 3270
QY 6264 atgttccagggcagatgacacagcalttgcctgggaagccagcttataccgggtgaaagaaatga 6323
|||
DB 3271 ATGTTCCAGCGCATGACACACATTTGGCTGGAGCCACACTTATCCCGGTAAACAACTGA 3330
QY 6324 tggctcggggaagaaacacatgcttcaatctcctaacttgccttgaaggaatgaagcttcaag 6383
|||
DB 3331 TGCTCGGGACAAAGAACCAATGCTTCAATCTGATCTGCTTTTAAAGCAATGATCTTACG 3390
QY 6384 ccggcactgtcttcttgcacaaatgtcgtgattacgttccggaagcgtgataccaccggttcc 6443
|||
DB 3391 CCGGCACTGTCTTTCGCAAAATGGCTGATTAAGTCGGAAAGGCTGATCCGACCCGGGTTTC 3450
QY 6444 agcaactatgaaggggtgagcccaacacggaagtttgaagcgcgaaccagatgtgaagcc 6503
|||
DB 3451 AGCACTATGAAGGGGTGACCCACAAACCGGAAGTTTGAAGCAGCGCACCAATTTGAAGGCC 3510
QY 6504 ggaatgatgtcgcggccaaagtaattgaagaatacttgaacaaatlaaacagcgaagcaat 6563
|||
DB 3511 GGAATGATGCTCCGGCCAAAGGTAATTTGAAGAAATACTTGAACAAATTAACCGCAAGCCAT 3570
QY 6564 ttactctgaattgaatacgtctcacgcagctatgggaactccgcttgaacgttgcgctcaca 6623
|||
DB 3571 TTATTCAGTTGAATAGCTACAGCCATGGGCAACTCCGTCGGTGAAGCTGGCGCTACA 3630
QY 6624 cggccttggaagaaataaccccaactacaaagggcgtctcaatctggagacttgaattgaaccaag 6683
|||
DB 3631 CGGCGCTGGAAGAAATACCCCACTACAGGCGGCTTCACTTGAGACTGGAATGACCAAG 3690
QY 6684 gacttgaaagaaagcgggcaactgtcttgaaggcgactcgaatgacggcgaagcagact 6743
|||||

D	3691	GACGCGAAAAAAGACGGGCACTCTTTATGCGGGGCGACCTTGATACGGGCCAACCGACT	3759
OY	6744	atgaattctcgcggaacgcgcgtgcttctgtgacccggaactgaaatcgcgaactgcta	6803
D	3751	ATGAAATTCTCGGGGAACGGCTGTGTTGGTCGACCGGACTGAATGCCGAACGCGCTA	3810
OY	6804	atgcaagggcccttaccgcaacttaagttaagaatgggaatggagcctcttcaca	6863
D	3811	ATGTCAAGGCCCTTTACGCCAACCTTAAGTTAGAAATAAATGATGGGACGCTCTTCACA	3870
OY	6864	aaaacgaattatattaccacaagctcaacttaactctctgactgactgcttggctcg	6923
D	3871	AAAACGCAATTTATTATTACCAACAGCTCATCTTACTCTTGTGACTAGCTTTTGGTGC	3930
OY	6924	atggcaagttgacctaccagaagccgagcctctgaccttggccttggagccttggcgatccg	6983
D	3931	ATGSCAAGTTGACCTTACACAGAGCGGCGCTGTGACCTTTGGCTGGAGCCTGGCAATCG	3990
OY	6984	ggaccttggcccttgccttggcgcggaagtgcgtgataaaaagggaggtcgctacgcgg	7043
D	3991	GGACCTTTGGCCCTGCGCTTGGCGGAAGTGCATGATGAAGAAAGAGGTCGTATCCGGG	4050
OY	7044	taacgagccaccttaaaagaagacttgccttggcgagatgaagagcttcaacttggatgaag	7103
D	4051	TTACGGCCCACTTAAGAGAACTTGTGGCGGATGAGGGCTTCACTGTGGCTGAG	4110
OY	7104	cagaagaagtacgtcacaagaacttgcgcggaatttaagccggaaggcgcgcaagattagtg	7163
D	4111	CAGAAAGATGAGCTCAAAACCTCGCGAATTTAAGCGGAGGGCGCAGATTATGTG	4170
OY	7164	attccgagctacaactcagagccttgaagaagataacttccaattctcttccaagtca	7223
D	4171	ATTTCCGACTCAACCTAAGGCTTAAAGGAAATAACTTCAAAATTTCTTCTTCAAGGTCA	4230
OY	7224	agggctgcgcgagttccctccaagtatgcgcgttagggaaataacttgaagcgcgctgcgagat	7283
D	4231	AGGGCTTGCCGGTTCCCTCAACTATGACCGGTAGGGAAATACTTGAAGCGGCTGCCGAT	4290
OY	7284	ttaacctctgcgcgccttgcagcaaacagcagccggaactgtttagcgctatgactcgcg	7343
D	4291	TTTACCTTCTGCGGGCGCCTGACGGACAAAGACGGGGAGCTGTTACGGCTATATCTGG	4350
OY	7344	cccggttggaagaatgcgcgcgaatgatgcgcgttgaagaacatcagcttgcgagtgcaag	7403
D	4351	CCCGGTGGGAAATGCGGCAATATGCCCCGTTGAAACATCAAGCTGCGAGGTCAAGG	4410
OY	7404	aaagctcgttttggccaagaagccttaagcttgccttgccttgaagggtaatttaa	7453
D	4411	AAGACTCGGTTTGGTCAMAGACTGCTTTACGTGTGCTGTCCCTTAAAGGGTGATTTAA	4470
OY	7464	ccgtgaccttaagaatcgatgtagacgcgggcgaagaattcgttlaacaactgacttccagcg	7523
D	4471	CCGTGACCTATGAAGTGGAGTGAAGGAGCGGGCAAGATTGGCTGTAAACACTGACTTCCAGCG	4530
OY	7524	cggagaagaagctggtctcttgcagccttggcttgaacctgagccctgcgcaagaagatga	7583
D	4531	CGGAGAGAGCGGCTACTTTCACGACCTTTGGCTGTGAACCTTACCCCTCGCAAAAGACTGA	4590
OY	7584	ccgaatccgctactctgtgtcgtgagacccaatagaagagcttaaccaaaccgctttggaagta	7643
D	4591	CCGATTACCCCTACTATAGCTGTGGACTTAATGAGAGCTTACCACCAACCGCTTGGAAAGSTA	4650
OY	7644	attaccctggacatcaccaagagcggtlaaaaaagaactttagccata--tctgcgcg	7700
D	4651	ATTACCTGGGCACTTACACAGGAGCGGTAAAAAAGAACTTTAGCCCATTAACCTGCTCCGC	4710
OY	7701	aggaacacgggcaaccggagagaaagttcgtgttaccagctctttagtgaagaagggcgagt	7760
D	4711	AGGAAAGGGCAACCGAGACGAAGTTCCGTGACACAGCTTTGTATGATAAAAAGGGCGGCT	4770
OY	7761	tgggaatttaagcccaatggtggcgagacttgaactgtctgcttgcacatctcgcgcc	7820
D	4771	TGGAAATTTACGGCAATATGGGCGAGACTTTAAGCTGTGCTGCTTTGGCAATATCTGCGGCC	4830

Oy	7821	aaatgaagcagcgacccacgcgttlttgactgacctacaataactctgggttagagcct	7880
Db	4831	AAATTTGAAGCACCAGGACCAACCGCTTTTGAACGTACTTAACAATTACCTTGGGTTAAGGCT	4890
Oy	7881	taagcgcccaagtggagggttcg9cg9ggaatgactcctctgg9g9caagaaggtccaaccg9aat	7940
Db	4891	TAAAGGCCCAAGTAGGGGGGTCCGGCGGGGATGCATCTCTGGGGGCGAGAAGGTCACCCCGAAT	4950
Oy	7941	tctgcctgatgctcaaaaagccgcagcgttcgccttgatgattcaagcccctttactaa	8000
Db	4951	TCTGCTGGATGCTCAAAAAGCCCCAGCTCCGCTGCTGATTCAGCCCTTTACTAA	5010
Oy	8001	aataa	8005
Db	5011	AATAA	5015
RESULT 4 EYE132038			
LOCUS	EYE132038	3840 bp	mRNA circular SYN 28-JUL-1999
DEFINITION	Expression vector pCDPT.		
ACCESSION	AJ132038		
VERSION	AJ132038.1 GI:5640088		
KEYWORDS	AMP gene; beta lactamase; ColE1 origin of replication; multiple cloning site; sp6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; xanthine-guanine phosphoribosyl transferase gene. Expression vector pCDPT. Expression vector pCDPT. artificial sequence; vectors. 1 (bases 1 to 3840) Zeng,B.J. Mammalian Expression Vector for fuse Xanthine-guanine phosphoribosyl transferase Tag unpublished 2 (bases 1 to 3840) Zeng,B.J. Direct Submission Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, 100080, CHINA		
SOURCE	ORGANISM		
REFERENCE	AUTHORS		
TITLE	JOURNAL		
JOURNAL	REFERENCE		
AUTHORS	TITLE		
TITLE	JOURNAL		
JOURNAL	FEATURES		
source	location/Qualifiers		
promoter	1..3840		
	/organism="Expression vector pCDPT"		
promoter	/db_xref="taxon:90749"		
	209..863		
promoter	/note="CMV"		
	864..882		
misc_feature	/note="T7"		
	882..984		
misc_feature	/note="Multiple Cloning site; HindIII, BamHI, BstXI, EcoRI, NotI, XhoI"		
	929..1387		
CDS	/codon_start=1		
	/product="Xanthine-guanine phosphoribosyl transferase"		
polyA_site	/protein_id="CA851567.1"		
	/db_xref="GI:5640088"		
promoter	/translation="MSRKYYVTMDMDIIRKRTASRUMPSEQMKGIIVSRGLYPCA LIALEELGIRHVDVICISSYHDNOREIKYLRAEGDEGFVIDDLVDTGTAVALRE MYPAHPVITFEAKPAGRPALDDYVDIDIPDITWIEOPWDMGVFVPISGR"		
	1649..1863		
rep_origin	/note="BGH"		
	2450..2775		
gene	/note="SP6"		
	2644..2729		
CDS	/note="SV40"		
	complement(2844..3704)		
CDS	/gene="amp"		
	complement(2844..3704)		
CDS	/gene="amp"		
	/codon_start=1		
CDS	/product="beta-lactamase"		
	/codon_start=1		

OY	888	gaacctggaagatgacacatccacatccatctcccttcctcaataaataatggaattgacgcga	947
Db	1230	gaccttggaagatgacacatccacatccatctcccttcctcaataaataatggaattgacgcga	1289
OY	948	ttgctctagtagagatgcatcattcattctctgagggatgaggtgaggaagacgaaggggga	1007
Db	1290	ttgctctagtagagatgcatcattcattctgagggatgaggtgaggaagacgaaggggga	1349
OY	1008	ggattggaagacaataagacgaatgctggga-	1040
Db	1350	ggattggaagacaataatagccgaatgacgcacacacgaccccaacgcccattcacgacg	1409
OY	1041	-----	1040
Db	1410	atttgcatgtccaccgccgcttctatgaaaggattggcgcttgcgaattgctttccggagc	1468
OY	1041	-----	1040
Db	1470	ccggctggaatgacgtctccacggcgggagatgcattgctggaagtcttccgcccaccgaact	1529
OY	1041	-----	1040
Db	1530	tggtttattgcagctttatgattggtacaaaataaagcaaatgacatfacaatattcacaaata	1589
OY	1041	-----	1040
Db	1590	aagcatTTTTTTTACATGCACTTCTAGTTGGTTTGTCCAAACTCAATGATGATCTATC	1649
OY	1041	-----	1040
Db	1650	atgtctgatatccgtgcacacctctagctagagcttggcgatfcatgctatgctagtctt	1709
OY	1041	-----	1040
Db	1710	ctgtgtgaattggttatccgcctacaaattccacacaaatagagcgcggaacatfaaagt	1769
OY	1041	-----tgcgatgggctctatg	1057
Db	1770	gtaaagcctgggggctctaataagtgatgacgtactacatttaattgggtggcctcactgc	1829
OY	1058	ctctct-----gaggggaagaagaccagctgcatatgaattgagcgaagcgag	1107
Db	1830	ccgctttccagtgccggaaacctgctctgcccagctgcatfataatgaaatcgccacacggcg	1889
OY	1108	ggaagagcggttttgcatlttgaggcgctcttcgctctccctcgcctcaactgactgcgtgcgt	1167
Db	1890	ggagaggggggttccgtaattggggcgcttccgcttccgcctcaactgactgactgcgtgcgt	1949
OY	1168	cgatcgcttcgagctgagcgagcggtatcagctcaactcaaaagcggtaatacgttatcca	1227
Db	1950	cggtctgttcggcttcggcgagcggtatcagctcaactcaaaagcggtaatacgttatcca	2009
OY	1228	cagaataaggagataaagcgaagaataatgagcaaaagccagcaaaagggccaga	1287
Db	2010	cagatcagggagataaagcgaagaataatgagcaaaagggccagcaaaagggccaga	2069
OY	1288	accgtaaagaagcgcgcttgatgagctttttccatagcttcgcccccttgacgagatc	1347
Db	2070	accgtaaagaagcgcgcttgatgagctttttccatagcttcgcccccttgacgagatc	2129
OY	1348	acaaanaatcgagctcaagtcagaagtgagcaaaaccgacagagactataaagatacag	1407
Db	2130	acaaanaatcgagctcaagtcagaagtgagcaaaaccgacagagactataaagatacag	2189
OY	1408	cgattcccccggaagtcacctgagctctctctcttcttgagacctgcccccttacggat	1467
Db	2190	cgattcccccggaagtcacctgagctctctctcttcttgagacctgcccccttacggat	2249
OY	1468	accgttcgcgcttctcctctccttcgggaagcgtagcttctcaatgactcaagctgtagt	1527
Db	2250	accgttcgcgcttctcctctccttcgggaagcgtagcttctcaatgactcaagctttagt	2309
OY	1528	atctcagttcggtagtagtctgctcgtccaagcttggtgcgtgtgtgcaaaccccccgctc	1587

Db	2310	ATTCAGTTCGGTAGTCCGTTGGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTC	2369
QY	1588	agcccgacccgctctgaccttaccgctaactcgttcttgaagtcacaccggttaagacagc	1647
Db	2370	AGCCCGAACCCGCTGGCGCTTATCCGSGTAACTATCGCTTGAFTCCAAACCCGGTTAAGACAG	2429
QY	1648	acttatcgcaactgycagcagcagccactgtttaacagatatacagagcgagatgtatgaagcg	1707
Db	2430	ACTTATCGCCCACTGGCAGCAGCACACTGGTAAACAGATTATGACAGCGAGCTATGTAGCGC	2489
QY	1708	gtgtacagaaattttttaagtgtgtggccctaactcagcctacacatagaaagacagattatgt	1767
Db	2490	GTGCTACAGAACTTTTGAAGTGTGGCCCTAACTACGGCTACACTAGAAAGACATATTGG	2549
QY	1768	gtatctgcgctctgctctaagcagcttaaccttcggaaaaaagatgtgtagcctctgataccg	1827
Db	2550	GTATCTGGCGCTCTGCTGAAGCGACCATTTACCTTCGGAAAAAGAGTTGTGTAGCTCTTGATCCG	2609
QY	1828	gcaaacaaacccacgcctgtgtagcggtgtgttttttltgtttgtaacagcagcagattcgcgca	1887
Db	2610	GCAAAACAACACCCGCTGGTAGCGGTGTGTTTGTGTTTGCACACAGACAGATTATGCGCGCA	2669
QY	1888	gaaaaaaggatctcagaagaatcctttgatactttctacaggggtctgacgctaagtga	1947
Db	2670	GAAAAAAGGATCTCAAGAAATGATCCTTTGATCTTTTTCACGGGGCTCTGACCGTAGTGCA	2729
QY	1948	acgaaactcagcttaagagatttgtgtcagatgagcagatacatatttgaatgtattaga	2007
Db	2730	ACGAAACTCACTGTTAAAGGATTTTGGTCATGACATTTATCAAAAAGATCTTTCACCTTAGA	2789
RESULT 6			
LPLPREP/c			
LOCUS	LPLPREP	2093 bp	DNA
DEFINITION	Plasmid pLP1 (from Lactobacillus plantarum CCM 1904) replication	BC1	24-Apr-1996
ACCESSION	protein pLP1 (rep) gene, complete cds.		
VERSION	M31223		
KEYWORDS	M31223.1 GI:149684		
SOURCE	replication protein.		
ORGANISM	Plasmid pLP1 (tissue library: CCM 1904) DNA.		
REFERENCE	1 (bases 1 to 2093)		
AUTHORS	Bouia,A., Brindel,F., Frey,L., Kammerer,B., Belarbi,A.,		
TITLE	Guyonvarch,A. and Hubert,J.C.		
JOURNAL	Structural organization of pLP1, a cryptic plasmid from		
MEDLINE	Lactobacillus plantarum CCM 1904		
COMMENT	Plasmid 22 (3), 185-192 (1989)		
FEATURES	Draft entry and computer-readable sequence for [1] kindly submitted		
Source	by B. Kammerer, 11-JAN-1990.		
	Location/Qualifiers		
	1..2093		
	/organism="Plasmid pLP1"		
	/plasmid="Plasmid pLP1"		
	/specific_host="Lactobacillus plantarum"		
	/db_xref="taxon:2593"		
	/tissue_id="CCM 1904"		
	249..1215		
	/gene="rep"		
	249..253		
	/gene="rep"		
	267..1215		
	/gene="rep"		
	/codon_start=1		
	/transl_table=11		
	/product="replication protein"		
	/protein_id="AA98164.1"		
	/db_xref="GI:149685"		
	/translation="MSIEFDKTEKNGKVRMRKRIENRYAYELALIEFKRHADRG		
	CGSVLVRKIGEHLLKLYQTFECHKRLCPCLNWRSSKNSNSLQIIIAEAAVRPKRF		
	LFTLTVVNAKSAEELKVS.LRALTLKAFNKLTRKATKVTKNLGLRSTELTIVSDSY		

```

repeat_L_region      NOHLHLVLFVKSSEYFNKSNNTYLAQAEAKIMOKALKVDYEPVHOAVKANKRKGTDS
                        LOASEETAKVEKSDADYMTADDERNLVVIKINLEYALAGTROIISYGILLKOIKODLKL
                        EDVENGDLVHVHGEDDYTRKEQHEAAEYVAKMDFKONVEIW"
1335..1335
repeat_L_region      /note="direct repeat A"
1352..1368
repeat_L_region      /note="direct repeat B"
1369..1385
repeat_L_region      /note="direct repeat C"
1386..1402
repeat_L_region      /note="direct repeat D"
1403..1419
repeat_L_region      /note="direct repeat E"
1420..1436
repeat_L_region      /note="direct repeat F"
1437..1453
repeat_L_region      /note="direct repeat G"
1456..1467
repeat_L_region      /note="direct repeat H"
1468..1484
repeat_L_region      /note="direct repeat I"
1485..1501
repeat_L_region      /note="direct repeat J"
1504..1520
repeat_L_region      /note="direct repeat K"
1521..1537
repeat_L_region      /note="direct repeat L"
1538..1554
repeat_L_region      /note="direct repeat M"
2051..2064
rep_origin            /note="origin of replication (put.): putative"
BASE COUNT           639 a 335 c 466 g 653 t
ORIGIN               1 bp upstream of HindIII site.

```

```

Query Match          15.3%; Score 1238.4; DB 3; Length 2093;
Best Local Similarity 98.0%; Pred. No. 2.6e-285;
Matches 1254; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```

```

QY 2032 gacacattcccccgaagaagcgacacgtacgacgagatcgagagatcatatccatgacat 2091
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1280 GCACAGCTGCACAAACAAAAGGATATGCGACAGGATTTACCGTTGATCATTCCTGACAT 1221

QY 2092 tctcttaccacaaataaatttggtttatataaatacccatlcttgagacacattctcc 2151
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1220 TCTCTTACCAATAAATAATTTGTTATTAATAATCCCATTTTGCACAACTTCTTCC 1161

QY 2152 gaagcttccatttgccttcttggtgtaattcattcgccacacatgaaataatcacattc 2211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1160 GCAGCTTCATTTGCTTTGGTGAATCTTCATCGCAACATGAACCTAATCACCATTTC 1101

QY 2212 tcaaatcttcaagtttcaaatcttggcttaatttggctttaataatcacatagctgatt 2271
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1100 TCAACATCTTCAAGTTTCAAAATCTTCTTAATTTGCTTAATATATCACCATATGCTGATT 1041

QY 2272 tgtcgtgtccagcttaaggacatactccaaattttaataccaccaaatacagctcatca 2331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1040 TGTGCGTTCAGGCTTAAGCATATCTCAAAATTTTAAATACCCCAAAATTAAGCCTCATCA 981

QY 2332 tcaagcgtacataataatcagctgatttaccctgtaattgcgcgttcttcgcgacatgct 2391
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 980 TCAGCGCTCATATTAACACCTGATTTTACCTCTATTTTGCCTTTCTTCGGGACATAGCT 921

QY 2392 tgcgaagagtcagttcctttagcttggtagctttaacagcgcgcacatgacacagcgc 2451
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 920 TGCAAAAGATGAGTTCCTTACGTTTGTTAGCTTTAAACGCTTGACATGACACACAGGCG 861

QY 2452 tcaataacattccaagccttlttgacataatttggccatcttgccttgcgttgtaataa 2511
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 860 TCATATATCAACTTCAAGGCTTTTTCGCAATATTTTGGCCATTCCTGCTGTGTAATAATA 801

QY 2512 ttattgaattcttaataaactgatttcaacaacagcaacacatgacagtgatgattat 2571
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 800 TTATTTGAATTTCTTAAATTAACCTGATTTTACAAACAGCACACATGCAAGTGTGATTA 741
QY 2572 tatgacccgtctgtgtcatatcaagcgttaattccgttgaacgttaataaataaataatt 2631
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 740 TATGACCCGCTCTGTTTCAATTAAGGTAATTTCCGTTGAACGTAATTAACCAATTAATTT 681

QY 2632 ttagtcaactttttatagcagtagcttatataaaggctttagttaaagcttcttaagac 2691
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 680 TTAGTACACTTTTATATAGGAGTATGCTTAATTAAGGCTTTTAAAGCTCTTTAAAGAC 621

QY 2692 actttaactctctgcctgaatgagcgtttttaacggttaaaagcttaaaaaaacaacgct 2751
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 620 ACTTTTAACCTCTCTGCTGAATGACGTTTAAAGGTTAAAGTTAAACAAAACCGT 561

QY 2752 cctttagcctctctgtgcaactgtcttcgcgaataatttgtttaactggctgagtttcc 2811
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 560 CCTTAAAGGCTCTGTTGAACACTGCTTCCGCAATTAATTTGTTTAACGTGCTGCAAGTTTTC 501

QY 2812 atgtcctcttccaaattacacaaatgagcaaatcgtttatgacaaaacacgcttgatata 2871
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 500 ATGCTCTCTTCCCAATTAACAAATGACACAAATCGTTTATGACAAAACACGCTTGATTA 441

QY 2872 agttttaagtgctgcacaaattcacaagaacgcaaaacttccacaaacccgatacatca 2931
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 440 AGTTTAAATGCTCTGCGCAATCTTACGAAAACCCAAACTTACACCAACCCGTTACATCA 381

QY 2932 tgtgccggtttaaattcctaagattgcgaataatctgcgaatagcgaacatttcaatctc 2991
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 380 TGTGCCCGGTTAAATTCATGAATGCGCAAAATATGCGCATGCGACATTTTCAATCTTTC 321

QY 2992 cgttcgcgaaggtcctaactttgcatttcaagtttatcttcaaaaaattcttaacata 3051
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 320 CGTTTCGCGCAAGGTTACTTATCTTGGCCATTTTCAAGTTTATTCACAAAATTTCTGACATA 261

QY 3052 aaaaagctctcagtttaccacacgtgaagaagcgtactactcttttcaataagcttata 3111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 260 AAAAGCTCTCCAGTTTATTCACGTAAGAGACTGACTATCTTTTTCATTAAGCTTATTA 201

QY 3112 ccttgacatcatagagcttccctcctagataagctataaactgcgaatgataaact 3171
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 200 CCTTGACATCATAGGCTTTTCCCTAGATAAGGCTATTAATTCGCAAAATGAATATCACT 141

QY 3172 caagtgcttcgcagcgcaaaactagaatttgcagctgagtttatttgccttcttct 3231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 140 CACGTTCTCCGAGCGGCCAACTAGGAATTTGCACGTGGGTTTATTTTGTCTTCTTT 81

QY 3232 caaccaatttaaaccttaataatacacaagaagcctataaaatcaatgatacaagccc 3291
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 CACGCAATTTTAAACCTTAATATACCAAAAAGCTTATTAATCAATGATGATCAAGGCC 21

QY 3292 aattaagcctaatacaagctt 3311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 20 AATTAAAGCCTAATCAAGCTT 1

```

```

RESULT 7
LOCUS      C300RFR      2140 bp      DNA      BCT      07-MAR-1996
DEFINITION Plasmid pc301l (from L. plantarum) ORF R, complete genome.
ACCESSION J03319
VERSION    J03319.1 GI:144180
KEYWORDS   complete genome.
SOURCE     plasmid pc301l DNA.
ORGANISM   plasmid pc301l
            plasmids.
REFERENCE  1 (bases 1 to 2140)
AUTHORS    Skaugen,M.
TITLE      The complete nucleotide sequence of a small cryptic plasmid from
            Lactobacillus plantarum
            plasmid 22 (Z), 175-179 (1989)
JOURNAL    MEDIANE
MEDLINE    90160873
COMMENT    Draft entry and computer-readable sequence for [1] kindly submitted
            by M.Skaugen, 19-NOV-1989.

```


JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters, H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelisstr. 5, D-24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
Location/Qualifiers
source 1..3986
/organism="synthetic construct"
/plasmid="pCDNA3.2EO"
/db_xref="taxon:32630"
misc_feature 1..2125
/note="cloning vector (pCDNA3) (Invitrogen)"
misc_feature 889..394
/note="multiple cloning site (MCS)"
misc_feature 2126..2796
/note="cloning vector (pZeoSV) (Invitrogen)"
misc_feature 2797..3986
/note="cloning vector (pCDNA3)"
BASE COUNT 900 a 1057 c 1054 g 975 t
ORIGIN

Query Match 13.3%; Score 1082; DB 56; Length 3986;
Best Local Similarity 100.0%; Pred. No. 7.3e-248;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatgaggggccaagatagcaggttgatgattatgacagttatataagatca 60
DB 210 gatgaggggccaagatagcaggttgatgattatgacagttatataagatca 269

QY 61 attacggggtcatatgattacatagccatataatgagatccgcgttacataactcgta 120
DB 270 attacggggtcatatgattacatagccatataatgagatccgcgttacataactcgta 329

QY 121 aatggcccgctgctgacacggccaagaccccgccattgaagtcataatgacgtat 180
DB 330 aatggcccgctgctgacacggccaagaccccgccattgaagtcataatgacgtat 389

QY 181 gtcccatagtaagcgaatagggacttcattagacgtcaatagtggtgacatttaagg 240
DB 390 gtcccatagtaagcgaatagggacttcattagacgtcaatagtggtgacatttaagg 449

QY 241 taactgcccacttgacagatcatcaatgtaataatgacagacgccccctatgac 300
DB 450 taactgcccacttgacagatcatcaatgtaataatgacagacgccccctatgac 509

QY 301 gtcaatgacgttaaatgacggcgcctgacattatgccagttacatgaactatggagctt 360
DB 510 gtcaatgacgttaaatgacggcgcctgacattatgccagttacatgaactatggagctt 569

QY 361 cctactggcagttacatagctatgatacgtatcgtattacatgagtggttgg 420
DB 570 cctactggcagttacatagctatgatacgtatcgtattacatgagtggttgg 629

QY 421 cagttacatcaatgagcggtgagatgagtggttgaactcaaggggattcccaagttccacccc 480
DB 630 cagttacatcaatgagcggtgagatgagtggttgaactcaaggggattcccaagttccacccc 689

QY 481 attgacgtcaatgaggtattgttttgacacaaatacaacggagacttcccaaatgctgt 540
DB 690 attgacgtcaatgaggtattgttttgacacaaatacaacggagacttcccaaatgctgt 749

QY 541 aagaactcggcccatgtagcgaatgagcggtgagtggttgaactgagtggtgaggttcatata 600
DB 750 aagaactcggcccatgtagcgaatgagcggtgagtggttgaactgagtggtgaggttcatata 809

QY 601 agcagagctctgctgacatagagaacccactgttactgagttacgaataatgaatcag 660
DB 810 agcagagctctgctgacatagagaacccactgttactgagttacgaataatgaatcag 869

QY 661 actcactatagagagaccgaagctgtgacgagctcgatccactagtaacgagcgcca 720

DB 870 ACTCACTATAGGAGAACCAACCTGTGATACGAGCTCGGATCCACTAGTAACGGCCGCCA 929

QY 721 gtgtgctggaattctgcagatatccatacactgagcgccgttgagatgattcaag 780

DB 930 gtgtgctggaattctgcagatatccatacactgagcgccgttgagatgattcaag 989

QY 781 ggcctattctatgtagtcaactaatgctagagctgcgtgacagctgagctgagct 840

DB 990 ggcctattctatgtagtcaactaatgctagagctgcgtgacagctgagctgagct 1049

QY 841 tctatgtccagacatcgtgtgttgccctcccccgtgaccttctgaacctggaagt 900

DB 1050 tctatgtccagacatcgtgtgttgccctcccccgtgaccttctgaacctggaagt 1109

QY 901 gccactccacgtctcttcccaataaataagagaattgtagctatgctgtgtagg 960

DB 1110 gccactccacgtctcttcccaataaataagagaattgtagctatgctgtgtagg 1169

QY 961 tgcattctatcttgagggtgagggtgaggcagagcaagggagagatgaggaagac 1020

DB 1170 tgcattctatcttgagggtgagggtgaggcagagcaagggagagatgaggaagac 1229

QY 1021 aatagcagcatcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1080

DB 1230 aatagcagcatcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1289

QY 1081 tg 1082

DB 1290 tg 1291

RESULT 9
AR098191
LOCUS AR098191 4026 bp DNA 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 4026)
AUTHORS Antelman, D., Gregory, R. J. and Willis, K. N.
TITLE Retinoblastoma fusion polyepitopes
JOURNAL Patent: US 6074850-A, 19 13-JUN-2000;
FEATURES
Location/Qualifiers
source 1..4026
/organism="unknown"
BASE COUNT 978 a 1020 c 983 g 1045 t
ORIGIN

Query Match 13.2%; Score 1068.6; DB 9; Length 4026;
Best Local Similarity 69.7%; Pred. No. 1.2e-244;
Matches 1918; Conservative 0; Mismatches 89; Indels 746; Gaps 5;

QY 1 gatgaggggccaagatagcaggttgatgattatgacagttatataagatca 60
DB 210 gatgaggggccaagatagcaggttgatgattatgacagttatataagatca 269

QY 61 attacggggtcatatgattacatagccatataatgagatccgcgttacataactcgta 120
DB 270 attacggggtcatatgattacatagccatataatgagatccgcgttacataactcgta 329

QY 121 aatggcccgctgctgacacggccaagaccccgccattgaagtcataatgacgtat 180
DB 330 aatggcccgctgctgacacggccaagaccccgccattgaagtcataatgacgtat 389

QY 181 gtcccatagtaagcgaatagggacttcattagacgtcaatagtggtgacatttaagg 240
DB 390 gtcccatagtaagcgaatagggacttcattagacgtcaatagtggtgacatttaagg 449

QY 241 taacatgccacttggcagtaacatcaagtgtatcatatgccaagtacccccatttgac 300
|||||
Db 450 TAAACTGCCACTTGGCAGTACATCAAGTGTATATATGCGAAGTACGCCCTATTGAC 509
QY 301 gtcaatgaagtaaaatggccgcctggcattatgcccagtaacatgaccttaaggactt 360
|||||
Db 510 GTCAATGACGGTAATATGCGCGCTGGCATATATGCCAGTACATGACCTTATGGAGCTTT 569
QY 361 cctacttggcagtaacatcctactgaatgtaagtcgtatatacctatgtaagtgattgg 420
|||||
Db 570 CTTCTTGGCAGTACATCTAGTATAGTCAATCCCTATTACCATGATGATCCGGTTTGG 629
QY 421 cagtaacatgaatggcgttggatagcgtttgactacaggggattccaagctccacccc 480
|||||
Db 630 CAGTACATCAATGGCGCTGGTATAGCGGTTGACTACAGGGGATTTCCAGTCTCCACCCC 689
QY 481 attgacatcaatggaggtttgttttggcaccacaatcaacgggacttccaatgtcgt 540
|||||
Db 690 ATTGACGTCATGGGAGTGTGTTTGGACCAAAATCAAGGGACTTTCCAAAATGTGCT 749
QY 541 aacaaatccgcctcattgaagcaaatggcggtgagcgttgaagtgagggtctata 600
|||||
Db 750 AACAACTCCGCCCTTACAGCCAAATGGCGGTAGCGGTGACGCTGAGCTATATA 809
QY 601 agcagaatcctctgcttaactagagaaccactgacttaactgactatcgaatataacg 660
|||||
Db 810 AGCAGAGCTCTCTGTGCTTACTAGAAACCCACTCTTACTGCTTATCAGAAATTAATACG 869
QY 661 actactataggagaaaccaagcttgg----- 687
|||||
Db 870 ACTACTATAGGAGACCCAAAGCTTTCGGCGGGTACCACTCTTCCGCATCGGTGCTG 929
QY 688 ----- 687
Db 930 CGAGGCGCACTGTGGGCTCGCGTTGAGAGCAAACTCTTCGGGCTTTCAGTACTC 989
QY 688 ----- 687
Db 990 TTGATGGAAGAACCCGTCGCGCTCCGAACGCTACTCCGCCACCGAGGAGCTGACGAGT 1049
QY 688 ----- 687
Db 1050 CCGCATGACCGGATCGGAAAACTCTCGAGCACTGAAAAACGAAAGTTAATGTGTA 1109
QY 688 ----- 687
Db 1110 AGTTAGTCTTTTGTCTTTTATTTATTTAGGTCCCGATCCGGTGTGTCAAATFCAAG 1169
QY 724 tgcgtgaattctgagatat----- 743
|||||
Db 1170 AACTGCTCTCAGTGTGATGTTGCCCTTACTCTTAGGCCCTGACGAAAGTGTACTTCTGC 1229
QY 744 ----- 744
Db 1230 TCTAAAAGCTGCGGAATTTGTAACCGCGCGCTGCACTAGAGCAATTCGCTGATGCA 1289
QY 775 ctgaagggccctattctatagtgtaacctaatagtctagagctcgctgatacagctcgact 834
|||||
Db 1290 TCGATGGGCCCTATTCTATAGTGTACCTAAATGCTAGAGCTCGCTGATCAGCTCGACT 1349
QY 835 gtgcctcttagttgcaagcaatcgtgtgttggccctcccccgtgcttcttcaacctg 894
|||||
Db 1350 GTGGCTTCTAGTTGCGAGCAATGTTGTTGGCCCTCCCGGCTCTCTTACACCTG 1409
QY 895 gaaggtgcacctccactgtccttctcctaataaatgagaaatgtacatgcatgtctg 954
|||||
Db 1410 GAAGGTGCACTCCCACTGTCTTCTTAATAAATGAGAAATGCAATTCGCTATGTCTG 1469
QY 955 agtagtgtaacttctactctgaggggtgaggttgggagcaagcaagggggaggtatg 1014
|||||
Db 1470 AGTAGGTGTCATCTATTTCTGAGGGGTGGGGTGAGGACAGCAAGGGGAGGATTTGG 1529
QY 1015 gaagacaataagcaggtcgtcgggga----- 1040

Db 1530 GAAGCATATACCGAAATAGACCAAGCAGCGCCCAACCTGCGATCAGCAGATTTTGA 1589
QY 1041 ----- 1040
Db 1590 TTCCACCGCCGCTTCTATGAAGAGTTGGGCTTCGAAATCGTTTTCCGGAGCGCGCTG 1649
QY 1041 ----- 1040
Db 1650 GATGATCTCCAGCGCGGGGATTCATGCTGAGTCTTTCGCCACCCCAACTTGTATT 1709
QY 1041 ----- 1040
Db 1710 TGCAGCTTATATGTTTACAAATAAAGCAATAGCATCACAATTTACAAATAAAGCAT 1769
QY 1041 ----- 1040
Db 1770 TTTTCACTGCATCTTAGTGTGTTGTGCCAACTCATGATATCTATCATGTCTG 1829
QY 1041 ----- 1040
Db 1830 TATACCGTCGACCTTAGCTAGAGCTTGGCGTAATCATGTATAGCTTTCTGTG 1889
QY 1041 ----- 1040
Db 1890 AAATTTTATCCGCTCACAAATTCACACACATACGAGCCCGGAAGCATTAAGCTGAAGC 1949
QY 1041 ----- 1040
QY 1041 ----- 1040
Db 1950 CTGGGTGCTCAATAGTAGTAACTCACTAATTAATTCGTTGGCTGCTACGTGCCGCTTT 2009
QY 1063 ----- 1114
Db 2010 CCAATCGGGAACCTGTGCTCCAGCTGCAATTAATGATCGCCAAACCGCGGAGAGG 2069
QY 1115 cgtttgtgtaattggcgctcttcgctctcgtactaactgactcgctcgctgctgct 1174
|||||
Db 2070 CGGTTTGGCTATTTGGGCGCTTTCGCTTCTCCTGCTCACTGCTGCTGCGCTCGTGT 2129
QY 1175 tgcgtcgcgcgcgcgtatcagctcactcaaaagcgtlaaactagcttaccacaagaatc 1234
|||||
Db 2130 TCGGCTGCGGAGCGGATACAGTCACTCAAGCGGTAATGAGTATCCACAGATC 2189
QY 1235 aggggataagcaggaagaacaatgtaagcaaaagccagcaaaagccaggaactaa 1294
|||||
Db 2190 AGGGATTAACCAAGAAAGCAATGTGAGCAAAAGGCCAGCAAAAGGCAGAAACCGTAA 2249
QY 1295 aaagccgcgttggctgagtttccataggtctcgcgcgccttgaagcaacaaana 1354
|||||
Db 2250 AAAGCCGCGTGGCTGGGTTTTCATAGGCTCCGCCCTCGACGATCACAAAAA 2309
QY 1355 tcaagctcaagtcagagtgtagcgaaccgcagagactaataagatacagagcgttcc 1414
|||||
Db 2310 TCGAGCGTCAAGTACAGAGTGGCGAAACCGACAGGATTAATAAGATACAGGGGTTTCC 2369
QY 1415 cccctggaagctccctcgttgcgtctcgttgcgaacctgcgcgttaccggaatacgttc 1474
|||||
Db 2370 CCGTGAAGCTCCCTCGGCTCTCTCTCCAGACCTCGCGTACCGGATACGCTGTC 2429
QY 1475 cgccttctcccttcgggaagcgttgcgttctccaatgctcaagctcgaagtgtaagtatcag 1534
|||||
Db 2430 CGGCTTCTCCCTTCCGGAACGTGGCGCTTCTCAATAGCTCAGCTGTAGCTATCTGAG 2489
QY 1535 ttcggtgaagtcgttcgctccaagcttgcgttgcgaagcccccgcgttcaagccga 1594
|||||
Db 2490 TTGCGTGAAGTCTGTGCTCCAAAGTGGGCTGTGTGACAGAACCCCGCTTCAAGCCGA 2549
QY 1595 ccgctgcgcttaccggttaactcgttgaatccaaccggtlaagaacgaactatc 1654
|||||
Db 2550 CCGCTGCGCTTATTCGCTGATCTGAGTCCAAACCGGTAAGACACGACTTATC 2609
QY 1655 gcaactgacagcaacacttgtaacaggaatlagagagcgaggtatgtagcgtgtcac 1714
|||||

D	b	1577	AGCTGGGGCTCGAGCATGCAAGCTTGAGTATTCATATAGTGTCACTAAATAAGTTGGCCT	1636
Q	y	1075	-----	1074
D	b	1637	AATCATGTCATAGCTGTTCGTGTGAAATTGTTATCCGCTCACAATTCCACACAA	1696
Q	y	1075	-----	1074
D	b	1697	TACGAGCCGGAGACATAAAGTGTAAAGCCTGGGGTGCCATATAGTGAGCTAACTACAT	1756
Q	y	1075	-----	1074
D	b	1757	TAAATGGGTTGGCTCACTCCCGCTTTCCAGTCGGGAAACTGTCGTGCGCAGCTGCATT	1816
Q	y	1087	aatgaatcgcgcaacgcgcgagggaagagcggttgcgatltgagcgcttctccgcttcc	1146
D	b	1817	AATGAATCGCGCAACGGCGGGGAGAGCGGGATTGGGTATTTGGGGCGCTTCCGCTTCT	1876
Q	y	1147	cgctcaatgactgcgtcgcgctggttcggtcgtcgctgcgagcggtatcaagctcaat	1206
D	b	1877	CGCTCACTGACTCGCTCGCTCGTGTTCGGCTGGGAGACGGGTATCAGCTCACTAA	1936
Q	y	1207	aggcggtaatcaggttataccacagaatcaggggataacgcaggaagaacatgtgacaa	1266
D	b	1937	AGCGGGTAAATACGGTTATCCACAGATTCAGGGGATTAACGAGAAAGAACTGTACGA	1996
Q	y	1267	aagcgacgcaaaagccaggaagccgtataaagcgcggttgcgttgcgcttltccatagc	1326
D	b	1997	AAGCGCAGCAAAAGCGCAGGAMCCGTAAAAAGCGCGCTGTCTGGCGTTTTCATAGGC	2056
Q	y	1327	tcgccccccctgcgaggaacatcacaaaaatcgcgcgtcgaagtacagatggtgcgaacccga	1386
D	b	2057	TCGCCCCCCTTGACGAGCATTCACAAAATCGAGCTCAAGTAGAGGTGGGAAACCCGA	2116
Q	y	1387	caggactataaataacacagcgcttcccccttggaagctccccctgctgcgtccctgtc	1446
D	b	2117	CAGACTATATAAGATACCAAGCGCTTTCGCCGTGAAGCTCCCTGCGCTCTCTCTGTC	2176
Q	y	1447	cgacccctgcgcttaccggaatacgttccgccttctcccttcggaagcgctgagcgctt	1506
D	b	2177	CGACCCCTGCGCTTACCGGATTCGCTTCGCGCTTTCCTTCGGGAAGCGTGGCGCTT	2236
Q	y	1507	ctcaatgtctaacgctgtgagatcatctcaagttcgtgtgagtgcttcgcgccaaagctgggct	1566
D	b	2237	CTCATACTCACTGATGTAGGTATCTCAGTTCGGTGTAGGTCTTGGCTTCGATCAAGCTGGGCT	2286
Q	y	1567	gtgtgcagaaacccccggttcaagcccgacgcgtgcgttalcggtgaatactgcgttgc	1626
D	b	2297	GTTGTGCAGAAACCCCGCTTCAACCGCAGCCGCGGCGCTTATTCGCGTAACTATCGCTTGG	2356
Q	y	1627	agtcacaacccggttaagaagaagacttatgcacactgcgacgacacactggttaacaggaata	1686
D	b	2357	AGTCCAAACCCGTTAAGACACAGACTTATTCGCACTGGCACACACCACTGTTAACAGATTA	2416
Q	y	1687	gcgaagcagagtatgtagcggtgcgtcacaagagttcttcaagtggtgcgtctaactcgcgtc	1746
D	b	2417	GCAAGGAGGATATGTAGGGGGGTATCAGAGATCTTTGAAGTGTGTGGCTTAATCAGGCT	2476
Q	y	1747	acactagaagaaacgatatattgtaatactgcgctctgcgtgaagccagtaacctcggaaaaa	1806
D	b	2477	ACACTAAGAGACAGTATTTTGTATCTGTGGCTCTGCTGAAGCAACTTACCTTCGGAATAA	2536
Q	y	1807	gaagtgttagctcttgaatcggnaaaacaaacacgcgtgtgtagcggtgttttttgc	1866
D	b	2537	GAGTTGTAGCTCTTGATTCGCGCAACAAACCAACCGCTGTGTAGCGGTGTTTTTTTGT	2596
Q	y	1867	gcaagcagcagattacgcgcagaaaaaagatctcaaaagaatcctltgactcttctta	1926
D	b	2597	GCAAGCAGCAGATTATGCGCGCAGAAAAAAGGAGATCTAATAAATCTTTGATCTTTCTTA	2656
Q	y	1927	cgaggctcgaagctcaagtgaacgaaaaactcaagttaaaggaatttggatcatgagcgat	1986

QY	1987	acataattgtaattagata	2007
Db	2717	CAAAAGGATCTTCACCTAGA	2737
QY	13		
LOCUS	AF053407	4050 bp	DNA
DEFINITION	Expression vector pMAS-S,		circular SYN
ACCESSION	AF053407		complete sequence.
VERSION	AF053407.1	GI:3746483	
KEYWORDS			
SOURCE			
ORGANISM	Expression vector pMAS-S.		
REFERENCE	Expression vector pMAS-S.		
AUTHORS	artificial sequence; vectors.		
TITLE	1 (bases 1 to 4050)		
JOURNAL	Krieg, A.M., Wu, T., Weeratna, R., Efler, S.M., Love-Homan, L., Yang, L.,		
MEDLINE	Yl, A.R., Short, D., and Davis, H.L.		
AUTHORS	Sequence motifs in adenoviral DNA block immune activation by		
TITLE	stimulatory CpG motifs		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)		
FEATURES	98454522		
SOURCE	2 (bases 1 to 4050)		
gene	Wu, T., Efler, S.M., Davis, H.L., Krieg, A.M. and Schorr, J.		
CDS	Direct Submission		
	Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725		
	Parkdale Ave., Ottawa, ON K1V 4E9, Canada		
	Location/Qualifiers		
	1..4050		
	/organism="Expression vector pMAS-S"		
	/db_xref="taxon:82305"		
	3351..4031		
	/gene="S"		
	3351..4031		
	/gene="S"		
	/note="derived from Hepatitis B virus surface antigen"		
	/codon_start=1		
	/transl_table=11		
	/product="surface antigen subtype ayw"		
	/protein_id="AAC64340.1"		
	/db_xref="GI:3746484"		
	/translation="MENITSGFGLPPLVLVDAGFFLLRLITITQSLDSMTSLNPLAG		
	TYVCLGNSQSPSTNSHSPTSCPPCPDGYRMCRLRIIFELILLCLIFLLVLVDG		
	MLPCLPIQSSTSTGCRCTMTAGTSMYSCCTKPSDNCNCIPIPSWAFGK		
	FLMMAASARFSLMLSLVPEVFMFVGLSPVWLIVIMMWYMSLISLSPFLPLPI		
	FFCLMYVI"		
BASE COUNT	993 a 1024 c 949 g 1084 t		
ORIGIN			
Query Match	11.9%; Score 968.6; DB 56; Length 4050;		
Best Local Similarity	85.9%; Pred. No. 1e-220;		
Matches 1155; Conservative	0; Mismatches 34; Indels 156; Gaps 1;		
QY	736	tgtcactcaaatgctagagctcgctgcatcagaacctgagactgtgcctctagtgtgcacgca	855
Db	40	TCTCGAGCTAGACTAGAGCTGCGTGATCGACCTCGACTGTGCTTGTAGTTGCCAGCA	99
QY	836	tctgttgtttgcgccctcccccgtgctctctccttgacacctggaagtgtgcacatcccatgtgc	915
Db	100	TCTGTTGTTTGCCCTCCCGTGCCTTTCCTTGACCCCTGGAAAGGTGCACATCCACATGTC	159
QY	916	ctttcctataaataagaaagaattgcatcgatgtgtcgaatgaatgtgtcattctatctg	975
Db	160	CTTTCCTATATAAAGAGAAATTCATCGCATTTGTCTGAGTAGGTGTCATTCATTCGTG	219
QY	976	ggggggtgggggtggggcaagcaagcaagggggaggattgggaagcaataagcaagcatgtc	1035
Db	220	GGGGGTGGGGGTGGGGCAGACAGCAGAGGGGAGGATTTGGGAAGCAATAGCAGGCATGCT	279
QY	1036	gggggagtcgggtgggctctatggc-----	1058

[illegible]

DB	1360	GTTAAGGATTTGGTCATGACGTG	1384
RESULT	14		
AF053408	AF053408	4227 bp	DNA circular SYN 17-OCT-1998
LOCUS	Expression vector pMCJ16-S,		complete sequence.
DEFINITION	AF053408		
ACCESSION	AF053408.1	GI:3746485	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
SOURCE			
gene			
CDS			
BASE COUNT	1021 a	1076 c	990 g 1140 t
ORIGIN			
Query Match	11.9%;	Score 968.6;	DB 56; Length 4227;
Best Local Similarity	85.9%;	Pred. No. 1e-220;	
Matches 1155;	Conservative 0;	Mismatches 34;	Indels 156; Gaps 1;
OY	736	tgtcaactaatgctagagctgcgtcatcgaatcgacctgacgtgtgaccttcaattgccagcca	855
DB	40	TCTCCAGTCTGACACAGAGCTCGCATGCAGCAGCTTCGACTTGCCATTGACAGCA	99
OY	856	tctgttgtttggccctccccccgctgcttccttgaccctggaagtgcaccccccacytc	915
DB	100	TCTGTGTGTTGCCCTCCCCCGTGCCTTCCTTAGACCCTGGAAAGTGCCACCTCCACTGTC	159
OY	916	cttctctaataaagtagaaattgcatcgcatgltcctaagtaagtgatcattcattctg	975
DB	160	CTTTCCTAATAAATGAGGAATTCATTCGATGTCGTGAGTAAGGTGTCATTATTTCTG	219
OY	976	ggggggtgggtggggcaggacagcaaggggagaatttgggaagacaataagcaggcatgct	1035
DB	220	GCGGCTGGGCTGGGCAAGCAAGCAAGGGGGAGGATTTGGGAAGAACAAATGACGAGCATGCT	279
OY	1036	ggggagtgagggtggtctcattgac	1058
DB	280	GGGGAAGGCTCGGACTGTGTGGCGTAATCATGTCATGTCATGTCGTGTGTAATTTGT	339
OY	1059		1058

```
Dh 340 TATCGCTCACAAATTCACACACATACGAGCCGGGAGAGCATAAAGTGAAGCGCTGG 399
Qy 1059 -----t 1059
Dh 400 GTGCCTAATGAGTACGTAACCTACATTAAATGCGTGGCTCACTGCCCGCTTCCAGT 459
Qy 1060 tctgaagcgaagaagaaccagctgcataatgaatcgcgaacgcgcggggaagcgcgtt 1119
Dh 460 CGGGAACCTCTCGTCCGACAGCTGAATGAATGCGCAACGCCGGGAGAGCGGCTT 519
Qy 1120 tgcgtatggcgctcttcgcgtcttcctgcctcgaactgcgtgcgtgcgtgcgtgc 1179
Dh 520 TGCCTATTTGGCGCTCTCCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 579
Qy 1180 tgcggcgagcgatgaactgaactgaactgaactgaactgaactgaactgaactgaact 1239
Dh 580 TGCGGCGAGCGGTATCAGCTCACTCAAGGCGTAAATACGTTATCCACAGAAATCAGGG 639
Qy 1240 ataacgcgagaagaacaatgtgaagcaaaagccagcaaaagccagcaaaagccgtaaaag 1299
Dh 640 ATAACGCGAGAAAGAACTGTGAGCAAAAGCCACAAAGGCCGAGAACCGTAAAGAG 699
Qy 1300 ccgcgtgcgtgcgttttccatagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1359
Dh 700 CCGGCTGCTGGCGTTTCCATAGGCTCCGCCCTCCTGACGAGCATCAGAAATCGAC 759
Qy 1360 gctaaagctcaagatgtgagcaaaagccagcaaaagccagcaaaagccgtaaaag 1419
Dh 760 GCTCAAGTCAAGAGGTGGGAAACCCGACAGACTAAAGATACAGCGCTTCCCGCTG 819
Qy 1420 gaagctccctcgtgcgtcctcctcgtcgcgaacctgcgcgtcgcgtcgcgtcgcgt 1479
Dh 820 GAAGCTCCCTGCTGGCGCTCTCCTGTTCCGACCTCCTGCGCTTAACGGATACGTCGGCT 879
Qy 1480 tctccctcggggaagcgtgcgcgtcttccatgcctcgaactgcgtgcgtgcgtgcgt 1539
Dh 880 TTCTCCCTCGGGAAGCGCTGGCGCTTCTCATAGCTCAAGCTGTAGTATCTCACTTCGG 939
Qy 1540 tctaggtcgtcgtcgtcgaagctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1599
Dh 940 TGTAGGTGTTGCTCCCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCT 999
Qy 1600 ggcgccttaaccgtaactatgcctcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1659
Dh 1000 GCGCTTATCGGTAACATATGCTGTGAGTCCACACCGGTAAGACACTTATGCGCAC 1059
Qy 1660 tgcgaagcgaactgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1719
Dh 1060 TGGCAGCAGCCACTGTGTAAACAGGATTACAGACGAGGATGTAGGCGGTCTACAGAT 1119
Qy 1720 tcttgaagtggtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1779
Dh 1120 TCTTGAATGTGTGCTTAACGCTACGCTACGTAAGAAAGATTTGTATCTTCGCTC 1179
Qy 1780 tgcgaagcgaactgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1839
Dh 1180 TGTGTAACCCAGTATACCTTCGGAAGAAAGTTGTAAGCTTGTGATCCGGCAACAAACA 1239
Qy 1840 ccgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 1899
Dh 1240 CCGCTGTGAGGCTGTGTTTTTTTGTTCAGACAGACAGATTACGGCGCAGAAAAAGAT 1299
Qy 1900 ctcaagaagatccttgccttctcgaagtgctgcagcgtcgaagtggaagcaaaatcac 1959
Dh 1300 CTCAAGAAGATCTTTGTATCTTTCTACGGGCTGTGACGCTAGTGAAGCAAAACTCAC 1359
Qy 1960 gttaaagatttgatgaagcgg 1984
Dh 1360 GTTAAGGATTTTGTATGAGCTG 1384
```

```
AF053409
LOCUS AF053409 4625 bp DNA circular SYN 17-OCT-1998
DEFINITION Expression vector pMC650-S, complete sequence.
ACCESSION AF053409
VERSION AF053409.1 GI:3746487
KEYWORDS
SOURCE
ORGANISM Expression vector pMC650-S.
REFERENCE Expression vector pMC650-S.
AUTHORS 1 (bases 1 to 4625).
Krieg,A.M., Wu,T., Weeratna,R., Efler,S.M., Love-Homan,L., Yang,L.,
Yl,A.K., Short,D. and Davis,H.L.
Sequence motifs in adenoviral DNA block immune activation by
stimulatory Cpg motifs
Proc. Natl. Acad. Sci. U.S.A. 95 (21), 12631-12636 (1998)
JOURNAL 9845422
MEDLINE 2 (bases 1 to 4625)
REFERENCE Wu,T., Efler,S.M., Davis,H.L., Krieg,A.M. and Schorr,J.
AUTHORS Direct Submission
TITLE Submitted (12-MAR-1998) HGD, Loeb Health Research Institute, 725
Parkdale Ave., Ottawa, ON K1Y 4E3, Canada
FEATURES
source
1..4625
/organism="Expression vector pMC650-S"
/db_xref="taxon:82307"
3926..4606
/gene="S"
3926..4606
/gene="S"
/note="derived from Hepatitis B virus surface antigen"
/codon_start=1
/transl_table=1
/product="surface antigen subtype ayw"
/protein_id="AAC64342.1"
/db_xref="GI:3746488"
/translation="MENVTSQFLGPIVLVLAQGFLLTRILTIPOSLSQSWTSLNPLCG
TTVCLGNSQSPHSPTSCPPPCPYRMWCLRRFLIFELLLCLFLVLLDLYG
MLPVCPLIPGSSITSPCRMTAGTSMPSCCTKPSDNGCTCIPSPSVAFK
FLMEMASAREFMSLSLVPFQWFEVGLSPYWLIVMMWVWGPSTLSLSPFLPLP
FECILMYVI"
BASE COUNT 1084 a 1193 c 1082 g 1266 t
ORIGIN
Query Match 11.9%; Score 968.6; DB 56; Length 4625;
Best Local Similarity 85.9%; Pred. No. 1e-220;
Matches 1155; Conservative 0; Mismatches 34; Indels 156; Gaps 1;
Qy 796 tgtcacctaagtctagctgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 855
Dh 40 TCTCGAGCTGAGACTAGAGCTGCTGATCAGCTGCTGACTGCTGCTGCTGCTGCTGCTG 99
Qy 866 tctgtgtttgcctcctcccggtgcctccttcgtacccctggaagtgccacccacgtc 915
Dh 100 TCTGTGTTTGGCCCTCCCGCTGCTTCTTGACCTGGAAGGTGCCACTCCCACTGTC 159
Qy 916 ctttcctataaagtgaagaatgcatgcgcatgtgcgaatggtgcgtcatctcattc 975
Dh 160 CTTTCTATTAATGAGGAATTCATGCAATGCTGTGAGTGTGATGATTTCTATTTCTG 219
Qy 976 gggggtgagggtgaggcagaagcagaagggaggatctggaagaacaatagacgacgtc 1035
Dh 220 GGGGCTGGGCTGGCGGACAGACAGCAAGGGGAGGATTTGGGAAGCAATAGCAGCATGCT 279
Qy 1036 ggggagtgcgtggggtcctatgac----- 1058
Dh 280 GGGGAAGGCTTCGACCTAGTGCGCTAATCATGTCATAGTGTCTGTGTAATTTGT 339
Qy 1059 ----- 1058
Dh 340 TATCGCTCACAAATTCACACACATACGAGCCGGGAGAGCATAAAGTGAAGCGCTGG 399
Qy 1059 -----t 1059
```

```
Db 400 GTGCTAATGAGTGAAGCTAACCTCAATTAAATGGCTTGCGCTCACTGCCCCCTTTCAGT 459
OY 1060 tctgaagcggaaaaaacacagctgcatctaataatcgcgaacacggcgaggagcggtt 1119
Db 460 CGGGAACCTCTCTGTCGCTGATTAATGAAATCGCAAGCGCGGGAGAGCGGGTT 519
OY 1120 tgcgtatcggcgcctctccctccctccctccctccctccctccctccctccctccctcc 1179
Db 520 TGGCTATTGGCGGCTCTTCCGCTTCCTCGCTACTGACTCGCTGCGCTCGGCTCGGCTCGGC 579
OY 1180 tgcggcgaagcggtatcagctcactcaaaagcggtaatacgggtatccacagaatcaagg 1239
Db 580 TGGCGGAGCGGATACAGCTCAAAAGCGGTAATACGGTATCCACAGAAATCAGGGG 639
OY 1240 ataagcagaagaagaacatgtagcaaaaagccagcaaaaagccagaacccgtataaaaag 1299
Db 640 ATAAAGCAAGAAAGAAATGAGCAAAAAGGCAAAAAGGCAAAAAGCCGTAAAAAGG 699
OY 1300 ccggttcgtgcgtttttccatagctccgccccctgaagacatcacaaaaatcgac 1359
Db 700 CCGGCTTGCCTGGCTTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAATCGAC 759
OY 1360 gctcaagtcagaagtgagcgaaccgacagactataaagaatacagcggtttcccccgt 1419
Db 760 GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCCCTG 819
OY 1420 gaagctccctcgtgcgtctccctcgttccgacctgcggttaacggatacctgtccgct 1479
Db 820 GAACCTCCCTCGCTCGCTCTCTCTCCGACCTGCGCGTACCGGATACCTGTCCGCT 879
OY 1480 ttcccttcggaagcgtgcgtcttctcaatgctcagctgtaggtatcagttcgg 1539
Db 880 TTCTCCCTTCGGGAAGCGTGGCGCTTCTCTATACGTCAAGCTGATGATCTCAATTGCG 939
OY 1540 tctagtcgttcgtcctcaagctggtgtgtcagcgaaccccggttcaagccgacgct 1599
Db 940 TGTAGGTGCTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGTTCAGGCCGACCGCT 999
OY 1600 ggcgcttaccggttaactatcgtcttgatccaaaccggttaagaacagactatcgccac 1659
Db 1000 GCGCTTATCCGGTAACTATCTGTAGTCAACCCGGTAAAGACAGACTTATCGCCAC 1059
OY 1660 tgcgaagcagccactggtgaacagagttagcagaagcaggtatgtagcggtgtctacaggt 1719
Db 1060 TGGCAAGCAAGCCACTGTTAACAAGATTAGCAAGCAGAGTATGTAGCGGTGCTACAGAGT 1119
OY 1720 tcttgaagtggtgccttaactacgctacactaagaagacagtaattggtatctgcgtc 1779
Db 1120 TCTTGAAGTGTGGCTTAACACTAGCGGTAACTAAGAACAGTATTGTATCTGCGCTC 1179
OY 1780 tgcgtgaagcagttactcctcgaaaaaagattgtagctcttgatccggcaacaacca 1839
Db 1180 TGCCTGAAGCCAGTTACCTTCGGAAAAAGATTGGTACTCTTGTATCCGGCAAAACAACA 1239
OY 1840 ccggttgtagcggtggtttttgttgcaagcagcagattacgcgagaaaaaaagat 1899
Db 1240 CCGCTGGTAGGGGTGTTTTTTTGTGTGCAAGCAGCAGATTACGCGCAGAAAAAAGGAT 1299
OY 1900 ctcaagaagatccttltgacttltctacgggtcgtcagtcagtggaacgaacactaac 1959
Db 1300 CTCNAAGAAATCTTTTATCTTTCTACGGGGTCTGACGCTCAGTGAACGAAAAACTCAC 1359
OY 1960 gtttaaggatttggatcatgagcg 1984
Db 1360 GTTAAAGGATTTGGTCATGAGCTG 1384
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 16:01:04 ; Search time 444.17 Seconds
(without alignments)
11471.782 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatctacggcgccagatatcac.....gcgtctgcctagcagcgtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: N_Geneseq_0601:*
2: /SIDSL/gcgdata/geneseq/geneseqn/NA1980.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn/NA1981.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn/NA1982.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn/NA1983.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn/NA1984.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn/NA1985.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn/NA1986.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn/NA1987.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn/NA1988.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3497.2	43.1	5059	21	AA236227
2	3110.6	38.3	3203	11	AA006760
3	1245.2	15.3	3853	19	AAV40006
4	1152.4	14.2	5306	21	AAAS3878
5	1148.2	14.1	14455	19	AAV32374
6	1148.2	14.1	14455	21	AAAS9050
7	1140	14.0	5377	21	AAAS3872
8	1131.4	13.9	4645	21	AAAS3870
9	1082	13.3	5446	19	AAV38297
10	1071.6	13.2	5245	21	AAAS3871
11	1070.2	13.2	4026	19	AAV40007

12	1068	13.2	8705	20	AA237778	Vector pShuttle DN
13	1058.4	13.0	6050	21	AAA47797	PMR9/SPO-S vaccine
14	1058.4	13.0	8002	21	AAA47796	PMR9/SPO-M vaccine
15	1036.2	12.8	3796	21	AAA27831	Vector plasmid PCM
16	988	12.2	4597	22	AAE24901	Nucleotide sequenc
17	984.6	12.1	8136	18	AAE63236	Plasmid PCR(TM)31a
18	981.8	12.1	6253	20	AA084454	AAV vector sequenc
19	981.8	12.1	6280	20	AA084455	AAV vector sequenc
20	981.8	12.1	6280	20	AA084456	AAV vector sequenc
21	981.8	12.1	6365	21	AA435921	Nucleotide sequenc
22	981.8	12.1	6981	21	AA45931	Nucleotide sequenc
23	981.8	12.1	7054	21	AA45927	Nucleotide sequenc
24	981.8	12.1	7405	21	AA45926	Nucleotide sequenc
25	973	12.0	4283	18	AAE64449	DNA encoding hgh.
26	973	12.0	4283	19	AAE6610	Epidermal growth f
27	973	12.0	4283	21	AAE63237	Plasmid PMR61630 n
28	961.4	11.8	6200	21	AA249493	Human alpha-L-idur
29	959.6	11.8	6115	18	AAE63235	Eukaryotic express
30	948.4	11.7	5857	21	AAE95423	Expression vector
31	935.6	11.5	5432	21	AAE89476	Transgenic APP DNA
32	932.8	11.5	5732	21	AAE71429	Expression vector
33	932	11.5	5729	21	AAE71430	Expression vector
34	921.8	11.4	6281	22	AAE84206	Plasmid Glambda-1A
35	921.8	11.4	6284	19	AAV41427	Plasmid Hu19AHcpd
36	921.8	11.4	6285	17	AAE34109	Anti-IL-5 humanise
37	921.8	11.4	6285	19	AAV03503	Plasmid pCD115H2HC
38	912.6	11.2	6094	21	AAE71428	Expression vector
39	908.2	11.2	5957	21	AAE55464	Destination vector
40	908.2	11.2	5957	21	AAE55467	Destination vector
41	906.4	11.2	2704	20	AAE06758	Vector pUC28 nucle
42	906.4	11.2	3130	19	AAV46334	Universal reporter
43	906.4	11.2	3364	21	AAE39495	Transgenic unc-119
44	906.4	11.2	3789	21	AAE39782	Plasmid pBIMMN nu
45	906.4	11.2	4173	19	AAV32663	Plasmid pGP-TT re

ALIGNMENTS

RESULT	1
AA236227	
ID	AA236227 standard; DNA; 5059 BP.
XX	
AC	AA236227;
XX	
DT	22-FEB-2000 (first entry)
XX	
DE	DNA encoding a permease an beta-galactosidase.
XX	
KW	Lactose operon; permease; beta-galactosidase; mutant; yoghurt;
KW	beta-galactosidase activity; lactose fermentation;
KW	fermented milk product; Streptococcus thermophilus; ds.
OS	
XX	Lactobacillus bulgaricus.
XX	
FT	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	122..1873
FT	/tag= a
FT	/product= "permease"
FT	/note= "encodes AA236220"
FT	1877..4519
FT	/tag= D
FT	/product= "beta-galactosidase"
FT	/note= "encodes AA236221"
XX	
XX	FR2778921-A1.
XX	
PN	26-NOV-1999.
XX	
PD	
XX	
XX	22-MAY-1998; 98PR-0006456.
PF	
XX	
XX	22-MAY-1998; 98PR-0006456.
PR	
XX	

(DANO-) CIE DANONE SA GERAIS.

Benbadis L, Brignon P, Gendreau F;
WPI; 2000-026146/03.
P-PADB: AAV53620, AAV53621.

New mutant *Lactobacillus bulgaricus* strain lacking beta-galactosidase activity useful for making fermented milk products, especially yoghurt

-

Disclosure: Page 15-22; 32pp; French.

The present sequence represents a lactose operon of *Lactobacillus bulgaricus*, and encodes a permease protein and beta-galactosidase protein. The specification describes mutant *L. bulgaricus* strain lacking beta-galactosidase activity, because it has a termination (nonsense) mutation in at least one of the coding sequences of the lactose operon. The resulting mutant strain is unable to ferment lactose. Since the mutant strain cannot ferment lactose, fermented milk products produced using it do not have to be cooled to prevent post-acidification. The mutant *L. bulgaricus* strain of the invention is used for producing a fermented milk product, especially yoghurt. Especially, milk is fermented with a culture of the mutant strain and optionally *Streptococcus thermophilus* in the presence of at least one sugar assimilable by the mutant strain, preferably glucose.

Sequence 5059 BP; 1229 A; 1341 C; 1323 G; 1166 T; 0 other;

Query Match	43.1%;	Score 3497.2;	DB 21;	Length 5059;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 3518;	Conservative	0;	Mismatches 13;	Indels 3;
			Gaps 1;	

OY	4584	ccgataaaaagttaagcgacggccttgcccttgctgctcagcagaagtgtaagctctacgcg	4643
Db	1479	ccgacacaaaagttaaggcgacggccttgcccttgctcagcagaagtgtaagctctacgcg	1538
OY	4644	cattggccgtactctccgacgcttgcgaagaaccgggacccgctgcctcgccctggaaatg	4703
Db	1539	cattggccgtactctccgacgcttgcgaagaaccgggacccgctgcctcgccctggaaatg	1598
OY	4704	aacatcgggctcttgctctgtatccacttgcgccttgcgcacggttcaatatcaaccgggactg	4763
Db	1599	aacatcgggctcttgctctgtatccacttgcgccttgcgcacggttcaatatcaaccgggactg	1658
OY	4764	gctctgtcaagctatgtttgaagaggcgacgaagtgaagccggccagcagatctcgtgaat	4823
Db	1659	gctctgtcaagctatgtttgaagaggcgacgaagtgaagccggccagcagatctcgtgaat	1718
OY	4824	tctgggaccggcgatctgaagcgagccaaagcttgaagaccggttaatcgtgacccgtcata	4883
Db	1719	tctgggaccggcgatctgaagcgagccaaagcttgaagaccggttaatcgtgacccgtcata	1778
OY	4884	acagcgaaacttccacaataatagcagatgctctcttcgcgaatcgccacacgctccaaagccc	4943
Db	1779	acagcgaaacttccgcacaatagcagatgctctcttcgcgaatcgccacacgctccaaagccc	1838
OY	4944	tggatgatgatctcaagttagaagagggaagaattagaaatgtgacaaataagttagttaaaag	5003
Db	1839	tggatgatgatctcaagttagaagagggaagaattagaaatgtgacaaataagttagttaaaag	1898
OY	5004	aaaaaagggttbaacaggcgacagccgagcccttgcttbaactbaaccgggaagttacggaagta	5063
Db	1899	aaaaaagggttbaacaggcgacagccgagcccttgcttbaactbaaccgggaagttacggaagta	1958
OY	5064	atacaattcccccgacctcgcagcaatagatcctctccaagaaccaggaagaagaactgtggagag	5123
Db	1959	atacaattcccccgacctcgcagcaatagatcctctccaagaaccaggaagaagaagaactgtggagag	2018
OY	5124	gcaagtcacagtttaatgtcagctcctctggacggggaactgtgctatgtactacgcttgaaaagc	5183
Db	2019	gcaagtcacagtttaatgtcagctcctctggacggggaactgtgctatgtactacgcttgaaaagc	2078

QY	5184	gccaaggagccagtcgaactctctatgccaagaagactcttgacgtacgtacaaatttaagtcatgca	5243
DB	2079	gccaaaggaccagctcaactctctatgccaagaagactcttacagatgaactttaagtcatgca	2138
QY	5244	aagtaacccggcaacccctggaaactccaagactcttgccagcccccagatgtgcacagctcaat	5303
DB	2139	aagtaacccggcaacccctggaaactccaagactcttgccagcccccagatgtgcacagctcaat	2198
QY	5304	atccatggaacgagcagtgaagagatcttcgccgcccccaaatctccaaagcaaaaatccgcctcg	5363
DB	2199	atccatggaacgagcagtgaagagatcttcgccgcccccaaatctccaaagcaaaaatccgcctcg	2258
QY	5364	ctctctatgtcaagatcttggaactcttgagaaagctcttcgggacaagaagatcagcttga	5423
DB	2259	ctctctatgtcaagatcttggaactcttgagaaagctcttcgggacaagaagatcagcttga	2318
QY	5424	agtttgaaaggggcggaacacgcacatcatgtccggttgaaacggccaactctgcctgcctcg	5483
DB	2319	agtttgaaaggggcggaacacgcacatcatgtccggttgaaacggccaactctgcctgcctcg	2378
QY	5484	gggaagactcctttaaccccaagcgaagttaatgtgtacccaagttcctccaagaagaata	5543
DB	2379	gggaagactcctttaaccccaagcgaagttaatgtgtacccaagttcctccaagaagaata	2438
QY	5544	accgcctgacggtgagcctcttaagaatattcttcgcgcctccctgcgtctggaagaccagagct	5603
DB	2439	accgcctgacggtgagcctcttaagaatattcttcgcgcctccctgcgtctggaagaccagagct	2498
QY	5604	ctcgtgcagatctccggtcttggtcttcagatcagtgacatcttcgaagccaaagccgcctgcacat	5663
DB	2499	ctcgtgcagatctccggtcttggtcttgagatcagtgacatcttcgaagccaaagccgcgtctgcacat	2558
QY	5664	tggaggaaccttaagcttaacggtccaagctctgaacccgatactacccaagaagaaagctcggaag	5723
DB	2559	tggaggaaccttaagcttaacggtccaagctctgaacccgatactacccaagaagaaagctcggaag	2618
QY	5724	tcgaagcccaatatgacctacgcgtcttgccaatgcccagctcttaagcttggaagtgcgggata	5783
DB	2619	tcgaagcccaatatgacctacgcgtcttgccaatgcccagctcttaagcttggaagtgcgggata	2678
QY	5784	gtgaaggtgactgtgtgtgtgaaaaagcttggtcccaatcagaagcgagccagcttggaattca	5843
DB	2679	gtgaaggtgactgtgtgtgtgaaaaagcttggtcccaatcagaagcgagccagcttggaattca	2738
QY	5844	ctctggtgctatttgccaggtagcgtgcgtctggggcgccggaagaagccttaaccttaacagagctc	5903
DB	2739	ctctggtgctatttgccaggtagcgtgcgtctggggcgccggaagaagccttaaccttaacagagctc	2798
QY	5904	gctctgattttaacaaaggcagcgagccactcttaaggtcttaagccggtgaagatgtggattcc	5963
DB	2799	gctctgattttaacaaaggcagcgagccactcttaaggtcttaagccggtgaagatgtggattcc	2858
QY	5964	gcaactcttgaaactaaagaacggtgataattgtactcttaacggtccagcggtatcttccaagg	6023
DB	2859	gcaactcttgaaactaaagaacggtgataattgtactcttaacggtccagcggtatcttccaagg	2918
QY	6024	gggccaacccggcagaaatttggaacgttaagtcttggtctgggtctatcccggaagaggtatga	6083
DB	2919	gggccaacccggcagaaatttggaacgttaagtcttggtccggtctatccacagaagaaggtatga	2978
QY	6084	tcctgggacatcaagaacatgaagcggaagcaacaatcaatgtctgtccgtgtctctaactacc	6143
DB	2979	tcctgggacatcaagaacatgaagcggaagcaacaatcaatgtctgtccgtgtctctaactacc	3038
QY	6144	cgaacacatctccctcttttaccggcctctgtgaagaagtaagcgtcttaagctcatgtatgaag	6203
DB	3039	cgaacacatctccctcttttaccggcctctgtgacaaagtaagcgtcttaagctcatgtatgaag	3098
QY	6204	ctaaactctgaaagccacggtcagctctgggaanaaagtgtgggggacagaaatcctatgctca	6263
DB	3099	ctaaactctgaaagccacggtcagctctgggaanaaagtgtgggggacagaaatcctatgctca	3158

QY	5604	tcvgcgacatctcggtttgtttcaagatcagttgactcttcaagccagccgctctgcact	5603s
Db	2499	tcctgscgacatctcgtgtttgtttcaagatcagttgactcttcaagccagccgctctgcact	2558s
QY	5664	tggaggaaccttaagcttaagccagcctgtgcagcatataccacaaaagaaagctgtgaag	5723s
Db	2559	tggaggaaccttaagcttaagccagcctgtgcagcatataccacaaaagaaagctgtgaag	2618s
QY	5724	tcgaagccaatatgtcctacccgtcttgccaatgvcagcctttaagctgtgaagtgcgggata	5783s
Db	2619	tcgaagccaatatgtcctacccgtcttgccaatgvcagcctttaagctgtgaagtgcgggata	2678s
QY	5784	gtgaaagtgacttggtgttcgtcgaaagctgtggcccaatcagaaagcgagccgctgggatatca	5843s
Db	2679	gtgaaagtgacttggtgttcgtcgaaagctgtggcccaatcagaaagcgagccgctgggatatca	2728s
QY	5844	ctctgcgtgactttgcagatgcgtgcctgcgagcgcggaagaaagcttaaccttaaccagctcc	5903s
Db	2739	ctctgcgtgactttgcagatgcgtgcctgcgagcgcggaagaaagcttaaccttaaccagctcc	2798s
QY	5904	gctcgttatattatcacagcgacgcctcttgaaggttaagccgagcaagatgtggtttcc	5963s
Db	2799	gctcgttatattatcacagcgacgcctcttgaaggttaagccgagcaagatgtggtttttcc	2858s
QY	5964	gcaactttgaaactaaagaagcggaattatgtaactcttaagccgagcggaatcgctttcaag	6023s
Db	2859	gcaactttgaaactaaagaagcggaattatgtaactcttaagccgagcggaatcgctttcaag	2918s
QY	6024	ggcgccaacccgacagaaatttgacagtaagtgttgctcgagctataccaggaagaggaatga	6083s
Db	2919	ggcgccaacccgacagaaatttgacagtaagtgttgctcgagctataccaggaagaggaatga	2978s
QY	6084	tctgtgacataaagaccaaataaagcggaagcaatcaaatgctgtccgtgcctctcaatcc	6143s
Db	2979	tctgtgacataaagaccaaataaagcggaagcaatcaaatgctgtccgtgcctctcaatcc	3038s
QY	6144	cgaaacccatccctcttttacccgctcgtgtacaagtaagcgcttaactgtcatgttgagag	6203s
Db	3039	cgaaacccatccctcttttacccgctcgtgtacaagtaagcgcttaactgtcatgttgagag	3098s
QY	6204	ctaactctgaaagccacagcgaactgtggaaanaagtgtggggcgagcaagatcctatgctca	6263s
Db	3099	ctaactctgaaagccacagcgaactgtggaaanaagtgtggggcgagcaagatcctatgcttca	3158s

QY 6264 atgtccagggcatgataccagcatgtgctgaggagccagtctatcccggtggaagaatgta 6323
|||||
Db 3159 atgtccagggcatgataccagcatgtgctgaggagccagtctatcccggtggaagaatgta 3218
QY 6524 tggctcgggacaaagaacatgcttcaatcctaactgtgctttaggcgaatgagctttaa 6383
|||||
Db 3219 tggctcgggacaaagaacatgcttcaatcctaactgtgctttaggcgaatgagctttaa 3278
QY 6384 ccggacatgtctcttggcccaaatggctgattacgtctccggaagagctgataccgacgggttc 6443
|||||
Db 3279 ccggacatgtctcttggcccaaatggctgattacgtctccggaagagctgataccgacgggttc 3338
QY 6444 agccctatgaaagggtgatacccaacacggaaagtttgagagcagccacccaagattgaaagcc 6503
|||||
Db 3339 agccctatgaaagggtgatacccaacacggaaagtttgagagcagccacccaagattgaaagcc 3398
QY 6504 ggaatgatagtctcggcgaaggttaattgaaagaatacttgaccaataaaaccaagccaagcat 6553
|||||
Db 3399 ggaatgatagtctcggcgaaggttaattgaaagaatacttgaccaataaaaccaagccaagcat 3458
QY 6564 ttatctcaagttgataagctacagctacagctatggcgaactccgtctgtgtgactgtgcgcttaca 6623
|||||
Db 3459 ttatctcaagttgataagctacagctacagctatggcgaactccgtctgtgtgactgtgcgcttaca 3518
QY 6624 cggccctcgaaataatacccccaactacacagggcggtcttcatctggaactgtgatactgacaaag 6683
|||||
Db 3519 cggccctcgaaataatacccccaactacacagggcggtcttcatctggaactgtgatactgacaaag 3578
QY 6684 gactggaataaagacggcgacactgtcttataggggcgacttcgaatgacggcgaacagact 6743
|||||
Db 3579 gactggaataaagacggcgacactgtcttataggggcgacttcgaatgacggcgaacagact 3638
QY 6744 atgaaattctcggggaagacggcgcttcttctgtaacgggaactcgacccaactgtgcta 6803
|||||
Db 3639 atgaaattctcggggaagacggcgcttcttctgtaacgggaactcgacccaactgtgcta 3698
QY 6804 atgcaagggcccttatacgccaacttaagttagaagtaaaagatgggcagactcttccca 6863
|||||
Db 3699 atgcaagggcccttatacgccaacttaagttagaagtaaaagatgggcagactcttccca 3788
QY 6864 aaaaagcaaatltaattacacaacagctcatcttactacttcttgactagcttcttggctcg 6923
|||||
Db 3759 aaaaagcaaatltaattacacaacagctcatcttactacttcttgactagcttcttggctcg 3818
QY 6924 atggcaagttgacctaccagagccggcgctctgacacttggcctgagagctgtgcgaatccg 6993
|||||
Db 3819 atggcaagttgacctaccagagccggcgctctgacacttggcctgagagctgtgcgaatccg 3878
QY 6984 ggaacttgccttgccttggccggaagtcgcgtgataaaaggaggagtcgtcttaccggg 7043
|||||
Db 3879 ggaacttgccttgccttggccggaagtcgcgtgataaaaggaggagtcgtcttaccggg 3938
QY 7044 taacggcccaacttaaaagaagacttgccttggcggaatgaggttcaactgtgctgaag 7103
|||||
Db 3939 taacggcccaacttaaaagaagacttgccttggcggaatgaggttcaactgtgctgaag 3998
QY 7104 cagaagaagtagctccaagaagctgcgggaatttaagccgggaagggcgccagattagttg 7163
|||||
Db 3999 cagaagaagtagctccaagaagctgcgggaatttaagccgggaagggcgccagattagttg 4058
QY 7164 attcggaactacaactagagcctgtaaaaggaataacttccaattctcttcccaagtca 7223
|||||
Db 4059 attcggaactacaactagagcctgtaaaaggaataacttccaattctcttcccaagtca 4118
QY 7224 agggctggccgggttccctcaagatagccggtaggaatacttgaacgggttcgggaat 7283
|||||
Db 4119 agggctggccgggttccctcaagatagccggtaggaatacttgaacgggttcgggaat 4178
QY 7284 ttaccttctggcggtccctgaagacaagcagccgggagagctggtttagcgtatgactgtg 7343
|||||
Db 4179 ttaccttctggcggtccctgaagacaagcagccgggagagctggtttagcgtatgactgtg 4238
QY 7344 ccggttggaataatgcgcgcaagtagccgcttgtaagaacatcagctgcgaggttcaag 7403

Db 4239 ccggttggaataatgcgcgcaagtagccgcttgtaagaacatcagctgcgaggttcaag 4298
|||||
QY 7404 aagactccgttttggctcaagaactcctttagctgtgctgtgctttaaagggtatttaa 7463
|||||
Db 4299 aagactccgttttggctcaagaactcctttagctgtgctgtgctttaaagggtatttaa 4358
QY 7464 ccgtgacctatgaagtgatgagcgggcaagattgtcgtatacaggtcgaactcccaagcg 7523
|||||
Db 4359 ctgtgacctatgaagtgatgagcgggcaagattgtcgtatacaggtcgaactcccaagcg 4418
QY 7524 cggagaagcgtgctctcttgcgaacgttggcttgaacctgacgttgcctgcgaagaactgta 7583
|||||
Db 4419 cggagaagcgtgctctcttgcgaacgttggcttgaacctgacgttgcctgcgaagaactgta 4478
QY 7584 ccgattacgcgtactatgctgtgtagacctaaatgtagagctacccaagccgcttgaaggtta 7643
|||||
Db 4479 ccgattacgcgtactatgctgtgtagacctaaatgtagagctacccaagccgcttgaaggtta 4538
QY 7644 attacgtgggatacttaccagggagcgttaaaaagaactttagcccata---tcgtccgc 7700
|||||
Db 4539 attacgtgggatacttaccagggagcgttaaaaagaactttagcccatactcgtccgc 4598
QY 7701 aggaacggggcaaccgggagcaaggttcgctgtgtaaccagctcttgatgaaaggcggt 7760
|||||
Db 4599 aggaacggggcaaccgggagcaaggttcgctgtgtaaccagctcttgatgaaaggcggt 4658
QY 7761 tgaatttacggccaatggggaagcttgaactgtctgtcttggccaatctgcgcgcc 7820
|||||
Db 4659 tgaatttacggccaatggggaagcttgaactgtctgtcttggccaatctgcgcgcc 4718
QY 7821 aaattgaaagcgggacccagcctttagaactgataacataacttgggttagagcct 7880
|||||
Db 4719 aaattgaaagcgggacccagcctttagaactgataacataacttgggttagagcct 4778
QY 7881 taagcgcccaatgaggtgctcgcggaatgtaactcctgtagggcagaaggttccaccgggaat 7940
|||||
Db 4779 taagcgcccaatgaggtgctcgcggaatgtaactcctgtagggcagaaggttccaccgggaat 4838
QY 7941 tctgcttgaatgctcaaaaagcccgccagcttgcgcgtgtagtcaagcccttactaa 8000
|||||
Db 4839 tctgcttgaatgctcaaaaagcccgccagcttgcgcgtgtagtcaagcccttactaa 4898
QY 8001 aataaagtctacaattgacttaacagatgaatttagtaaaagcaaaagagtaggga 8060
|||||
Db 4899 aataaagtctacaattgacttaacagatgaatttagtaaaagcaaaagagtaggga 4958
QY 8061 agatggcaacgatacagaagaatgccaagggcagccggcggtgcgtactagcgaaggt 8114
|||||
Db 4959 agatggcaacgatacagaagaatgccaagggcagccggcggtgcgtactagcgaaggt 5012

RESULT 2
AAQ06760
ID AAQ06760 standard; DNA: 3203 BP.
XX
AC AAQ06760;
XX
DT 25-FEB-1991 (first entry)
XX
DE Sequence of beta-galactosidase gene.
XX
KM PBR329; pBG1; ds.
XX
OS Lactobacillus bulgaricus SBR0034.
XX
FH Key Location/Qualifiers
FT CDS 285..3083
FT /*tag= a
PN JF02261383-A.
XX
PD 24-OCR-1990.

QY 6719 gacttcgatgaccgagccaaaccgactatgaattcttcgcgaggaacggcctgtcttgctgac 6778
|||||
Db 1797 gacttcgatgaccgagccaaaccgactatgaattcttcgcgaggaacggcctgtcttgctgac 1856
QY 6779 cgcagctgaatcgcgcgaacactgcttaattgaaggcccttaacgcgaaccttaagttagaa 6838
|||||
Db 1857 cgcagctgaatcgcgcgaacactgcttaattgaaggcccttaacgcgaaccttaagttagaa 1916
QY 6839 gtaaaagatgggagcagctctctcctcaaaaacgacaatttattcaaccaagctcatcttac 6898
|||||
Db 1917 gtaaaagatgggagcagctctctcctcaaaaacgacaatttattcaaccaagctcatcttac 1976
QY 6899 tacttttactagatcttttctgctgaatgaggaattgacctaccagagccgagcctctgacc 6958
|||||
Db 1977 tacttttactagatctttctgctgaatgaggaattgacctaccagagccgagcctctgacc 2036
QY 6959 ttctgacctgagagccttgcgaaatccggagaccttctgaccttgccgcggaagtcgctgat 7018
|||||
Db 2037 ttctgacctgagagccttgcgaaatccggagaccttctgaccttgccgcggaagtcgctgat 2096
QY 7019 gaaaaagggagagtcgctctaaccgggtacacggcccaacttaagaagaacttgctggcg 7078
|||||
Db 2097 gaaaaagggagagtcgctctaaccgggtacacggcccaacttaagaagaacttgctggcg 2156
QY 7079 gatgagggcttcaactgtgctgagagcaagaagtagcttaagaagctgcgggaatttaag 7138
|||||
Db 2157 gatgagggcttcaactgtgctgagagcaagaagtagcttaagaagctgcgggaatttaag 2216
QY 7139 ccggaagggcgcgccaatatagttgatttcgcgaatacaacctagaagcctgaagaataaac 7198
|||||
Db 2217 ccggaagggcgcgccaatatagttgatttcgcgaatacaacctagaagcctgaagaataaac 2276
QY 7199 ttccaaattcctctctccaaagtcgaagggtcgcggttccctcaagtatgcggttagg 7258
|||||
Db 2277 ttccaaattcctctctccaaagtcgaagggtcgcggttccctcaagtatgcggttagg 2336
QY 7259 gaatacttgaaggggtgcgcggaatttactcttcgcgccgaccttgacggaacacggccgg 7318
|||||
Db 2337 gaatacttgaaggggtgcgcggaatttactcttcgcgccgaccttgacggaacacggccgg 2396
QY 7319 ggaagcgtgtacgcgcgatgatctgcgcgggtgggaaatcccgcaagtatgcgccttg 7378
|||||
Db 2397 ggaagcgtgtacgcgcgatgatctgcgcgggtgggaaatcccgcaagtatgcgccttg 2456
QY 7379 aaagacatcagctgcgaggtlcaaggaagactccggttltggtcaagaactgcctttaggt 7438
|||||
Db 2457 aaagacatcagctgcgaggtlcaaggaagactccggttltggtcaagaactgcctttaggt 2516
QY 7439 cctgtgccttaaggggtgatttaaccgtgactatgaagtctgagacggcgcaagatt 7498
|||||
Db 2517 cctgtgccttaaggggtgatttaaccgtgactatgaagtctgagacggcgcaagatt 2576
QY 7499 gctgtacagcgtgacttcccaaggcggaagaagctgctctcttcgcaagccttgcttg 7558
|||||
Db 2577 gctgtacagcgtgacttcccaaggcggaagaagctgctctcttcgcaagccttgcttg 2636
QY 7559 aaactgagcctgcgcaaaaagactgacagattacgcgtactatgltctggaaccttagag 7618
|||||
Db 2637 aaactgagcctgcgcaaaaagactgacagattacgcgtactatgltctggaaccttagag 2696
QY 7619 agctaccaggaacgcgtctggaaggtaattctcctggtgcatctaccagagagcggtataaaaag 7678
|||||
Db 2697 agctaccaggaacgcgtctggaaggtaattctcctggtgcatctaccagagagcggtataaaaag 2756
QY 7679 aactttagccata--tctgcgcagagaaacggcaacccggagcaaggtctgctgtac 7735
|||||
Db 2757 aactttagccataactgtctgcgcagagaaacggcgcaacccggagcaaggtctgctgtac 2816
QY 7736 cagctctttgatgaaaaagggcggttggaatttaacggccaatggggcagacttgaaactg 7795
|||||
Db 2817 cagctctttgatgaaaaagggcggttggaatttaacggccaatggggcagacttgaaactg 2876

QY 7796 tctgcttgcatactctgcgcgcccgaattgaaagcagcggaaccagcgtcttgaactgact 7855
|||||
Db 2877 tctgcttgcatactctgcgcgcccgaattgaaagcagcggaaccagcgtcttgaactgact 2936
QY 7856 aacaatcaccttgggttaagccttaagcgcgccaagatgggggtcggcgggatgactcc 7915
|||||
Db 2937 aacaatcaccttgggttaagccttaagcgcgccaagatgggggtcggcgggatgactcc 2996
QY 7916 tgggggcagaaggtccaccgggaattctgcttgatggtccaaaaagcccgccagcttgc 7975
|||||
Db 2997 tgggggcagaaggtccaccgggaattctgcttgatggtccaaaaagcccgccagcttgc 3056
QY 7976 ctggtgattagcccttcttaaaaaataatgctacaattgacttaacagatgaaat 8035
|||||
Db 3057 ctggtgattagcccttcttaaaaaataatgctacaattgacttaacagatgaaat 3116
QY 8036 ttagtaaaagcaaaagcagtagaggaagatgcaacgatacgaagagtgccaagcgacgcg 8095
|||||
Db 3117 ttagtaaaagcaaaagcagtagaggaagatgcaacgatacgaagagtgccaagcgacgcg 3176
QY 8096 gcgtgtcgtcagcgagcgt 8114
|||||
Db 3177 gcgtgtcgtcagcgagcgt 3195
RESULT 3
AAV40006
ID AAV40006 standard; DNA; 3853 BP.
XX
AC AAV40006;
XX
DT 15-FEB-1999 (first entry)
XX
DE plasmid pCTM.
XX
XX
KW E2F: transcription factor; human; retinoblastoma protein RB;
bladder cancer; restenosis; angioplasty; diabetic retinopathy;
thyroid hyperplasia; Grave's disease; psoriasis;
benign prostatic hypertrophy; Li-Fraumeni syndrome;
peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
OS Chimeric - cytomegalovirus.
OS Chimeric - mastadenovirus.
OS Chimeric - bacteriophage T7.
OS Chimeric - bacteriophage SP6.
OS Chimeric - rhesus macaque polyoma virus.
OS Chimeric - Bos taurus.
XX
XX
FT Key Location/Qualifiers
FT promoter 209..864
FT /tag= a
FT /note= "CMV promoter"
FT m1sc_feature 907..1131
FT /tag= b
FT /function= tripartite leader sequence
FT promoter 1132..1149
FT /tag= c
FT /note= "SP6 promoter"
FT m1sc_feature 1679..3853
FT /tag= d
FT /note= "pUC19 backbone H3 to AatLI"
FT CDS complement (2857..3717)
FT /tag= e
FT /note= "AMP-ORF"
XX
PN MO9821228-A1.
PD 22-MAY-1998.
XX
PD 13-NOV-1997; 97WO-US21821.
XX
PF 14-FEB-1997; 97US-0801092.
XX
PR 15-NOV-1996; 96US-0751517.

[illegible]

XX	Synthetic.
OS	
XX	WO2000049162-A2.
PN	
XX	
XX	24-AUG-2000.
PD	
XX	
PE	22-FEB-2000; 2000WO-US04429.
XX	
PR	19-FEB-1999; 99US-0253022.
PR	08-MAR-1999; 99US-0263814.
PR	26-MAR-1999; 99US-0276820.
XX	
PA	(ATHE-) ATHERSYS INC.
XX	
PI	Harrington JJ, Sherf B, Rundlett S;
XX	
DR	WPI; 2000-549276/50.
XX	
PT	Non-targeted activation of endogenous genes, e.g. for the production of erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT	proteins and for drug discovery
XX	
DS	Disclosure; Fig 35; 240pp; English.

XX
CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDs), preferably the vectors
CC comprise (in sequential order) a TRS, an USDs, a rare cutting
CC restriction site (RCSR) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.

XQ Sequence 5306 BP; 1303 A; 1365 C; 1383 G; 1255 T; 0 other;

Query Match	14.28;	Score 1152.4;	DB 21;	Length 5306;
Best Local Similarity	76.48;	Pred. No. 8.2e-164;		
Matches 1524;	Conservative	0;	Mismatches 396;	Indels 74; Gaps 6

QY	23	gttaacattgattatgtaactagttatcaatgaatcaatcagggtgcaattgattcata	82
Db	140	gttgcaattgattctatttctactcagttatttaacagtaacaaatgaagggttcattgattcata	199
QY	83	gccacataatgaggttcgcggttaacataactaacgtgaanaatggccgcctgctgacgcgc	142
Db	200	gcccatatgtgagttccgcggttaacataacttaacgtgaanaatggccgcctgctgacgcgc	259
QY	143	ccaaacgccccccgcctcttgaactgaatgaatgaatgactgtgtcccatagtgaacgccaatag	202
Db	260	ccaaacgccccccgcctcttgaactgaatgaatgaatgactgtgtcccatagtgaacgccaatag	319
QY	203	ggaattcccaattgacgcgaatgggtggaactattacagtaaacctggccactctggcaatc	262
Db	320	ggacttcccaattgacgcgaatgggtggaactattacagtaaacctggccactctggcaatc	379
QY	263	atcaagtgtatcatabatgccaagtacagcccccattgacgtcgaatgacgtgaanaatggccgc	322
Db	380	atcaagtgtatcatabatgccaagtacagcccccattgacgtcgaatgacgtgaanaatggccgc	439
QY	323	ccctggcaattatgcccaatgaatgaactcttaagggaacttccctactctggcagaataatctacg	382
Db	440	ccctggcaattatgcccaatgaatgaactcttaagggaacttccctactctggcagaataatctacg	499

```
QY 383 tatatgcatcgtattaccatggtgatgcggttttggcagttacatgagcgctgagat 442
Db 500 tattatgcatcgtattaccatggtgatgcggttttggcagttacatgagcgctgagat 559
QY 443 agcggttctgcatcagcggtgatttccaaagtctccaccacattgagctcaatggagtttgt 502
Db 560 agcgcttctgcatcagcggtgatttccaaagtctccaccacattgagctcaatggagtttgt 619
QY 503 tttagcacaataatcaagcgagcttccaaaatgtcgttaacacat-----ccggcccc 554
Db 620 tttagcacaataatcaagcgagcttccaaaatgtcgttaacacatgcatgcgcgcgcgc 679
QY 555 attagcacaataatggcggttagcggtgtacaggttgggaggttctatatagcagagctctcgt 614
Db 680 gtttagcacaataatggcggttagcggtgtacaggttgggaggttctatatagcagagctctcgt 739
QY 615 gctactatgagaacccactgcttactatgcttatcgaaattaaacactacataagga 674
Db 740 agtgaacccgttagatcactgtgaattctgacgaactactgtattaaagcgcatagagctctcc 799
QY 675 gaccacaagctgtgacagcgagctcgatccactagtaacgagcgcgagctgtgtg- gaatt 733
Db 800 ttgcagaaatgtcttaagtacacactatcgatttccacacattatacagagccgagtgttaatt 859
QY 734 ctgacagatatacacaacttgcgcgcgcgtcgagcatgcatctagagggccctattctat 793
Db 860 gtcaacagcatcatgatacgtctcccgagagcagacaagccccaacatgctcgtcagtaata 919
QY 794 agtctcaactataatgcttagagctcgtctatcaagctcgacgtgtgctctcctagtccagc 853
Db 920 cgaactcaataatagggcgagcagtgtagtactcgtcaacttaagagag----- 965
QY 854 catctgttcttgcgcgcgcgcgcgcgcgccttcccttgcacccctgaaaggtgcacactccacatg 913
Db 966 -----gcatctcgtcgacaggttaccgcttgcagcgttcgaagaaagaagt 1000
QY 914 tcccttccataaataatgagaatttgcatactgctgtctgagtgagtgtaattctatcc 973
Db 1001 ttaagagagcgccaagaacagcgtcatgagccgaagtgcgagcccgatcttcccatcgt 1060
QY 974 tggggggtgggttggggagagacagcgaggggagagatttgggaagacatatagcagctg 1033
Db 1061 gtgactgtcgtgcatatagtcgcacgaacccgaacctgtgcgcgcgtgactgcgcgcacg 1120
QY 1034 ctggggatgcgctggtgctctatgcttgcagggcggaagaacccgcgcataatgatt 1093
Db 1121 atggc-----tcgcgcgttagagatccacagaacgggtgtgtgtgcacatga 1166
QY 1094 cggccaacgcgcgggagagagcggttgcgtatttggcgctcttcgcgtccctccgtccac 1153
Db 1167 tgcgctagctgatatgtgtctcaagtagcgaagcagacgtggcgcgccgaagc 1226
QY 1154 tgaactcgtgctgcgtgcgttgcgtctgcgcgcgcgcgcgcgttalcactactcaaaagcggt 1213
Db 1227 ggtcggagagtgcttcgcg-agaacgggtgtgcataagaaattcatcaacgcatatgacgt 1285
QY 1214 aatcgcgttatcccaacagatcaggggatacgaacgaggaagaacagatgttagcaaaagcgca 1273
Db 1286 agatccctgtcagtagtctgagatcgtctgagc-----catgttagcaaaagcgca 1334
QY 1274 gcaaaagcgcaagacgttaaaagggcgcgcttgcgtgcttttccataagctccgcgc 1333
Db 1335 gcaaaagcgcaagacgttaaaagggcgcgcttgcgtgcttttccataagctccgcgc 1394
QY 1334 cccctgacgagatatacaaaaatcgacgtctcaagtgacaggtgtgcgaaacccgcagagact 1393
Db 1395 cccctgacgagatatacaaaaatcgacgtctcaagtgacaggtgtgcgaaacccgcagagact 1454
QY 1394 ataaagataccagcgcttcccccgtgaagctccctcgtgcgcgtctcctgtttccgaacct 1453
Db 1455 ataaagataccagcgcttcccccgtgaagctccctcgtgcgcgtctcctgtttccgaacct 1514
QY 1454 gccgcttaccggtatcactgttcgcgccttctcccttcgggaagcgttgcgcttctcaatg 1513
```

```
Db 1515 ggcgcttaccggtatccctgtccgccttctcccttcgcggaaagcgttgcgcttctcatag 1574
QY 1514 ctcaacgctgtaagtattctcaagttcgtgtgaggttcgttcgctcaagctgtgctgtgtga 1573
Db 1575 ctcaacgctgtaagtattctcaagttcgtgtgaggttcgttcgctcaagctgtgctgtgtga 1634
QY 1574 cgaacccccgcttcaagccgcagcgtctgcgccttaccgcgttaactatcgcttgagtcaca 1633
Db 1635 cgaacccccgcttcaagccgcagcgtctgcgccttaccgcgttaactatcgcttgagtcaca 1694
QY 1634 cccggtlaagacacagactatcgccacttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1693
Db 1695 cccggtlaagacacagactatcgccacttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1754
QY 1694 gaggatgtgagcggtgtcgtacacagattcttgaagtggtgggcttaactacgcttaacgag 1753
Db 1755 gaggatgtgagcggtgtcgtacacagattcttgaagtggtgggcttaactacgcttaacgag 1814
QY 1754 aaggacagatattgttatctgcgtcgtcgttgaagccagttaccttcggaaaaagattg 1813
Db 1815 aaggacagatattgttatctgcgtcgtcgttgaagccagttaccttcggaaaaagattg 1874
QY 1814 tagctcttgatccggccaacaaacacacgcgctgtgtagcgtgttcttctgttgcgaagca 1873
Db 1875 tagctcttgatccggccaacaaacacacgcgctgtgtagcgtgttcttctgttgcgaagca 1934
QY 1874 ggaatttcgcgcgaaaaaaagataatcaagaatccttgatcttcttcaaggggc 1933
Db 1935 gcaagtaacgcgcaaaaaaagataatcaagaatccttgatcttcttcaaggggc 1994
QY 1934 tgacgtcagtggaacggaatactcacgttaaggaatttgtcatgagcggaatacat 1993
Db 1995 tgacgtcagtggaacggaatactcacgttaaggaatttgtcatgagcggaatacat 2054
QY 1994 tgaatgtattaga 2007
Db 2055 gatcttcacctaga 2068

RESULT 5
AAV32374/C
ID AAV32374 standard; DNA; 14455 BP.
XX
AC AAV32374;
XX
DT 29-SEP-1998 (first entry)
XX
DE Complete sequence of the pEI/Fiber plasmid.
XX
KW Circular; adenovirus type 5; pEI/Fiber plasmid; structural protein;
KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;
KW early gene; Huntington's disease; Tay-Sachs disease;
KW sickle cell disease; E1 regulatory protein; ds.
XX
OS Synthetic.
XX
XX Key location/Qualifiers
FT misc_feature 1460..4998
FT
FT /tag= a
FT /note= "AD5 E1 regulatory gene"
FT misc_feature complement (10922..14223)
FT
FT /tag= b
FT /note= "AD5 fiber gene consisting of a CMV
FT promoter at 5' end of this gene"
XX
XX PD 02-APR-1998.
XX PF 24-SEP-1997; 97WO-EP05251.
XX PR 25-SEP-1996; 96US-0719806.
```

```
XX (NOV ) NOVARTIS AG.
PA (SCRI ) SCRIPPS RES INST.
PI Memerow GR, Von Seggern DJ;
XX WPI; 1998-230709/20.
XX
PT Adenoviral vectors - which lack DNA encoding for structural protein
PT or fibre protein used particularly for gene therapy
XX
XX Example 1; Pages 112-131; 170pp; English.
XX
CC The present sequence is that of a pE1/fiber plasmid used in the
CC method of the invention. The plasmid contains an adenovirus type 5 (Ad5)
CC fiber gene controlled by a CMV promoter, an Ad5 E1 gene and a pMAM
CC backbone. The invention provides adenoviral vectors having deletions of
CC all or part of various gene sequences encoding adenoviral structural
CC proteins and/or early region proteins. Deletions in these proteins would
CC allow a reduced risk of wild-type virus contamination and would also
CC allow packaging of foreign DNA in such vectors for a variety of
CC diagnostic and therapeutic applications. The adenoviral vectors having
CC deletions in the structural and/or early gene regions are produced by
CC cellular complementation of these adenoviral genes. Therefore, the
CC pE1/fiber plasmid was used as a complementation plasmid which was
CC introduced into a host cell line where parts of the fiber and E1 gene
CC region would be stably inserted into the host cell chromosomes. The
CC resulting E1/fiber gene deficient plasmid can be used as a gene delivery
CC vector. The vectors can be used for diagnosis or gene therapy, e.g. for
CC treating conditions characterised by hyper-proliferative cells (e.g. for
CC tumours), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease,
CC or sickle cell disease), or infections (e.g. HIV infection). They can
CC also be used for in vitro production of biologically active proteins.
XX
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;
XX
Query Match 14.1%; Score 1148.2; DB 19; Length 14455;
Best Local Similarity 95.5%; Pred. No. 3e-163;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
XX
QY 755 ggcggccgtcagatgcatactagaggccctattctatagttcaccctaagtctagag 814
DB 11236 GCGGCGGCTCGACATGATCTAGAGGCCCTTATTCTATAGTGTACCTTAAATGCTAGAG 11177
QY 815 ctgcgtcatcagcctcgtactgtccttctagttgcagccatcgttgttgcctccc 874
DB 11176 CTGCGTGTAGTCAGCTCGACTGCTCTGTAGTTGCCAGGCATCGTGTGGCCCTCC 11117
QY 875 ccgtgcttcccttgagccctggaagtgtgccactccactgtccttctctaataatgag 934
DB 11116 CCGTGCCTCTCTTACCCCTGGAAGGTGCCACTCCACTGCTCTTCTAATAAATGAGG 11057
QY 935 aaattcatcgcactgtctgaftaggtgttcattctattcgggggggtgggttggcag 994
DB 11056 AAATTGCAATGCAATTGTCTGAGTAGGTGTATCTATTCTGGGGGGGGGGGTGGGAGG 10997
QY 995 acagcaaggggggaattgggaagacaatagcagcactgtctgggagatcggttggccta 1054
DB 10996 ACAGCAAGGGGGAAGATTGGGAAGACATAGCAGCATGCTGGGGAGTGGGCTGCTA 10937
QY 1055 tggcttcctgaagcgaagaagacagctgcatatgaatcggccaacgcggggagag 1114
DB 10936 TGGCTTCTGAGGGCGTA-----TGCGTGTGAATAACCCACAGATGCGTAAAGAG 10886
QY 1115 cggttgcgtatctgggcgtctcctcgtctcgaactcgaactgactggtgcgcgt 1174
DB 10885 AAATATCCGCAATAGGCGCTTTCGCTCTCTGCTACTGACTCGTGGCTGGCTGT 10826
QY 1175 tcggcttcgagcagatcatcactcaacaaaggcgtlaataacgttaccagaatc 1234
DB 10825 TCGGTCGGCGGCGAGCGGTATCAGCTCACTCAAGGCGGTAAATACGGTTATCCACAGATC 10766
```

```
QY 1235 aggggataacgcaggaagaacatgttgaagcaaaagccagcaaaaggccaggaacgtaa 1294
DB 10765 AGGGGATTAACGCAGGAAGAACAATGTGAGCAAAAGGCCAGCAAAAGGCCAGAACCTTAA 10706
QY 1295 aaagccgcgttgccttgccgttttccataggtccgcgcccttcagcagcatcacaanaa 1354
DB 10705 AAAGGCCGCGTGTGCTGGCGTATTTCATAGGCTCCGCCGCCCTCAGCAGCATCACAANA 10646
QY 1355 tcgaagctcaagtcagaggtgtggcgaaccgcagcagcatataaagataccagcgcttc 1414
DB 10645 TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACAGCGGCTTCC 10586
QY 1415 cccctggaagctccctgtgtgagctcctgtttccgaaccttcggttacggataccgttc 1474
DB 10585 CCCTGGAGGCTCCCTCGTGGCGCTCTCGTTCGACACCTCGGCTTACCGGATACCTGTG 10526
QY 1475 cgccttctcccttcgggaagcgtgagccttccaaatgtcctcaagcgttgaatcctcag 1534
DB 10525 CGCCTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAGCTCAGCGTGTAGTATCTCAG 10466
QY 1535 ttcggttgaagtgcttgcctccaagctgtgctgtgtgcaagaccgccgttcaagccga 1594
DB 10465 TTCGCTGTAGGTGCTGCTGCCAAGCTGGGCTGTGTCCAGCAACCCCGTTACGCCGA 10406
QY 1595 ccgctgagccttaccggttaactatcgtcttgaagccaacccggttaagaacgactatc 1654
DB 10405 CCGCTGCGCTTATCCGTAACATATGCTTGAAGTCCAAACCCGTAAGACACGACTTATC 10346
QY 1655 gcaactgcagcaacccactgtgaacagatatacagagcgaagatgtagcgtgtctac 1714
DB 10345 GCCACTGCGACGACGACCACTGTGAACAGAGATTAGCAGAGCAAGATATAGCGGCTGTAC 10286
QY 1715 agattcttgaagtgtgtgctctaactacggtctacataagaagacaglatttgtatctg 1774
DB 10285 AGATTCTTAAAGTGTGTGGCTGACCTACACTACACTAAGAACAGATATTTGGTATCTG 10226
QY 1775 cgcctgtgaagcaagttaccttcgcgaaaaaggttggtagccttgatccggcaaca 1834
DB 10225 CGCTCTCTGAAGCAAGTACCTTTCGGAANAAGTGGTAGTCTTGTGATCCGGCAACA 10166
QY 1835 aaccacgcgtgtgagcgtgtgtttttgtttgttgcgaagcagacagatlaacgcgagaanaa 1894
DB 10165 AACCAACGCGTGTAGCGGTGTTTGTGTTTGTGCAACAGCAGATTTACGCCACAAAAA 10106
QY 1895 aggatctcaagaagatccttgaatcttcttaacggygtctgaagcgtcagtggaagaaa 1954
DB 10105 AGGATCTCAAGAAGATCTTGTGATCTTTCATACGGGCTGTGACGCTCAGTGGAAAGAAA 10046
QY 1955 ctcaagttaaaggatlttgtaagagcgatataatlttgaaatgtattaga 2007
DB 10045 CTCACGTTAAAGGATTTTGTGTCATGAGATTATCAAAAAAGATCTTACCTAGA 9993
XX
RESULT 6
AAA59050/C
ID AAA59050 standard; DNA; 14455 BP.
XX
XX AAA59050;
AC
XX
XX 07-NOV-2000 (first entry)
DE
XX Nucleotide sequence of the E1/fiber-expressing plasmid pE1/fiber.
XX Adenovirus: tripartite leader; adenovirus vector particle; gene delivery;
KM E1 gene; fiber gene; ss.
XX
XX Synthetic.
OS Adenovirus type 5.
XX
XX WO200042208-A1.
XX
XX PD 20-JUL-2000.
XX
```


PF 14-JAN-2000; 2000WO-EP00265.
XX
PR 14-JAN-1999; 99US-0115920.
XX
PA (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERN GES MBH.
PA (SCRI) SCRIPPS RES INST.
XX
PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
PI Skripchenko Y;
XX
DR WPI: 2000-4/6068/41.
XX
PT New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy -
XX
PS Example 1; Page 160-164; 212pp; English.
XX
XX The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL
CC exons. The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC quiescent adenoviral vector particle. The present sequence represents
CC a El/fiber-expressing plasmid, which was used for complementation of
CC El/fiber-gene deleted adenoviruses.
XX
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;

Query Match 14.1%; Score 1148.2; DB 21; Length 14455;
Best Local Similarity 95.5%; Pred. No. 3e-163;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
QY 755 gggcgccgtcagcatgcatctagagggccctatctatctagtgtaacctaaatgctagag 814
DB 11236 GCGGCGCGTCAGCATGATCTAGAGGGCCCATTTCTATAGTGCATCAATAGCTAGAG 11177
QY 815 ctgcgtgcatcagcctcgacgtgctctctagttgccaagccatgctgttgcctccccc 874
DB 11176 CTCCTGTATCAGCCTCGACTGCTGCTTCTAGTTGCCACGCACTGTGTGTCCTCCCTCC 11117
QY 875 ccgtgacctcttgacctggaaggtgacactccactgctcttccctaataaagtgg 934
DB 11116 CCGTGCCTTCTTACCTCGAAGGTGCCACTCCACGTCTTCTTCTATATAATGAAG 11057
QY 935 aaatgcatcgcattgtctgtagtgtagtctatctctggtgggtggtggtggtggtggtg 994
DB 11056 AAATTGCAATGCGATTTGTGAGTAGTGTCTATCTATCTGGGGGGTGGGGGGGCAAG 10997
QY 995 acagcaaggaggagatgtggaaagacaatagcaagcagctggtgggtggtggtggtg 1054
DB 10996 ACAGCAAGGGGAGGATGTGGAAAGACATAGCAGCATGTGGGATGCGGGGCTCTA 10937
QY 1055 tggcttctgaaagcgaaagacccagctgcatatgaatcgcccaagcggtgggaaggg 1114
DB 10936 TGGCTTCTGAAGCGGTA-----TGGCGGTGAATAATCCGCACAAATGCGTAAGAG 10886
QY 1115 cggttggtcttggtggtggtcttccgtctctgctcagctgagctgagctgagctggt 1174
DB 10885 AAATATACCGCATCAGGCGCTCTTCGCTTCTCCTGCTCATGCTGCTGCGGTGCT 10826
QY 1175 tggcgctgagcgagcgatcatcagctcaaaagcggttaatacggatcatccagaatc 1234
DB 10825 TCGGCTGGGAGGAGGATATCAGCTCACTCAAAAGCGGTAATATACGTAATC 10766
QY 1235 aggggataacgaggaagaacatgtgagcaaaagcgagcaaaagcgaggaacggtaaa 1294
DB 10765 AGGGATTAACGACGAGAAAGACATGTGAGCAAAAGGCCACGCAAAAGCGACGAGAACCGTA 10706

QY 1295 aaagcgcgctgtcgtgagcttcttccatagctccgcccccctgacagacatcaaaaaa 1354
DB 10705 AAAGCCCGGTTGCTGCGGCTTTTCCATAGGCTCGGCCCTCGAGAGCATCACAAAAA 10646
QY 1355 tgcagctcaagtcagaagtggtggaaccgacaggaactataaagatacagcgcttcc 1414
DB 10645 TCGAGCTCAAGTCAGAGATGCGGAACCCGACAGACTATTAATACACAGCGCTTTC 10586
QY 1415 ccctggaagctccctcgtgagctctcgtcttcagaccctgacagcttaacgatactg 1474
DB 10585 CCTGGAAGCTCCTGTCGCTGCTCTCTGTCACCTGCGCTTACCGGATFACCTGTG 10526
QY 1475 cgctctcctccctcgaggaaagcggtgctcttctcaatgctacagcgctgtagtctcag 1534
DB 10525 CGCCTTTCCTCCCTTGGGGAAGCGTGGCGCTTCTCATAGCTCAGCGTGAAGTATCTAG 10466
QY 1535 ttcggtgtagtgcgttcgctcgaagctgggtgtgtgcaagaaaccccgctcagccga 1594
DB 10465 TTCGGTGAAGTGTGCTTCGCTCAAGCTGGGCTGTGTGCAAGAACCCCGCTTCAGCCGA 10406
QY 1595 ccgtgcgcccttaccgtaactatcgtcttgagtcacaaccggtgaagacgacttacc 1654
DB 10405 CCGTGCCTTATCCGGTACTATGCTTGAAGTCCAGCCGTAAGACAGACTTATC 10346
QY 1655 gccactggcagcagccactggttaacagagtagcagagcgagtagttagcggtctac 1714
DB 10345 GGCATGTGGAGCAGCAGCATGTAACAGAGTAGCAGAGGAGTAGTATGCGGCTAC 10286
QY 1715 agagttctgaaagtgtgtgcttaactaagcgctaacactagaagaaagatattgtgactg 1774
DB 10285 AGAGTCTTGAAGTGTGTGCTTACACTACGCTACACTGAAGAGAGATATTTGGTATCG 10226
QY 1775 cgctgtgctgaaagcagtaactctcggaaaagagtttgtagctctgtagtccggcaaca 1834
DB 10225 CGCTCTGCTGAAGCCAGTATACCTTCGAAAAAGAGTTGTTGATCTTGAATCCGCAACA 10166
QY 1835 aaccacgcgtgtgtagcggtgttcttctgttcgaagcagcagatlaacgagcaaaaaa 1894
DB 10165 AACCAACCGGTGTACCGGGGTTTTTTTGTGTGCAACGACGAGATTAACGCCGAAAAA 10106
QY 1895 aggatctcaagaagatccttgaatcttctcaggggtctgagcgtcagtggaagaaaaa 1954
DB 10105 AGGATCTCAAGAAAGATCTTTTATCTTTTCTACGGGGTCTGACGCTCACTGGAAGAAA 10046
QY 1955 ctcaagttgaaggtatttgatcagtgagcgagtaacatatttgatattaga 2007
DB 10045 CTCACGTTAAGGAGATTGTGTCATGAGATTATCAAAAGATCTTCACCTAGA 9993
RESULT 7
ID AAA53872 standard; DNA: 5377 BP.
XX
AC AAA53872;
XX
DT 03-JAN-2001 (first entry)
XX
DE Expression vector pRIG14.
XX
KW Vector; endogenous gene; activation; over-expression;
KW erythropoietin; growth hormone; drug discovery;
KW granulocyte colony stimulating factor; ds.
OS Synthetic.
PN WO200049162-A2.
XX
PD 24-AUG-2000.
XX
PF 22-FEB-2000; 2000WO-US04429.
XX
PR 19-FEB-1999; 99US-0253022.
PR 08-MAR-1999; 99US-0263814.

OY	615	gcaactatagagaaccacatctcttaactgtgctatctcgaaataatagacatccactatagga	674
Db	739	agtgaaacggtcagatcacctctgaatctgac---gacctacgtataacggtcatalagagcc	795
OY	675	gacccaagctctgtgacccgagactcgatcacatagataagccgcgcagctgtgtctggaaatcc	734
Db	766	tcctcgcagatcaacacagaagctttattctggttaattatcaacgtaataatgtcatacgag	855
OY	735	tgcagatatccatcaacacctgcgcgcgtctcgacatgcatctagagagccctattctata	794
Db	856	tcagctgccttcgacaacaacagctcgcgaacttaagctgcagtgactctcttaataccacca	915
OY	795	gtgtcacactaaatgtctagagctcgctgcatgaagctcgactgtgtgctctctagtccagcc	854
Db	916	tgtgctaacaggtgtagtactcgtctacctaagaagagagcctactctgcgcaggttaagcagtcgaa	975
OY	855	atctgtgtttgtgccctcccccgcgtgcctctcccttgaaacctggaaggtgagccacacagct	914
Db	976	gaaagaagtttaagaagagccggaataacagcgctcatagaccggaagtgtgcgagccggaatc	1033
OY	915	cccttccataaanaatgaggaanaatgtgcatgcgcatctgtctgagtaagttgtcatctattct	974
Db	1036	tcaccatcgtgtgatctgcgcgatatatagcgcgcagcaacgcacactgtgtgcgc-----	1088
OY	975	gggggggtgtgggtgtgggcagagacagcgaaagggtgagagatctgggaagacataagcagcatgc	1033
Db	1089	-----ggctgagctgcgcgcgcacagatctgcgtccgcggttagagatccacaagacggtgtc	1133
OY	1035	tggggatgacggttgggtctctatgtctctcttgagggcggaagaacaggtctgataatgaatc	1094
Db	1140	gtgtccgcattgt-----cgagctagctcgataagttgc	1166
OY	1095	ggccacagcgcggggagagagcggttctgtcgtattgtggcgtctctccgcctccctgcactcat	1154
Db	1169	ctccacactaaggagaaagcagacgacatctgtgggtggcgcgcaagaagcgtgtcgagcagtgctccga	1228
OY	1155	gactcgcgtcgcgtctgtcgtctgcgtgcgtgcgcgcagcgtatctcagctcaactcaaaagcggtata	1214
Db	1229	gaagcgggttcgcgcatagaagaattgtcatacagcatatagcgttagatctctgtctagaagtcga	1288
OY	1215	atagcgttatccacaagaatcagagggtataacgcgcgagaaagaacatgtgtgacaaagcgcaag	1274
Db	1289	gattcgtgtcgagc-----catgtgagcaaaagcgcaag	1319
OY	1275	caaaagcgcagagaaacgctaaagaagccgcgtgtcgtgcgtctttccatagagctccgcgcc	1334
Db	1320	caaaagcgcagagaaacgctaaagaagccgcgtgtcgtgcgtctttccatagagctccgcgcc	1379
OY	1335	ctgtacgagcatctcaanaaatactgacgcgtctcagctcagaaggttggcgnaaaccgcgcagactata	1394
Db	1380	ctgtacgagcatctcaanaaatactgacgcgtctcagctcagaaggttggcgnaaaccgcgcagactata	1433
OY	1335	taagaataccagcgcttcccccctgtgagaagctccctctgtgcgtctctccctgttccgaacctg	1454
Db	1440	taagaataccagcgcttcccccctgtgagaagctccctctgtgcgtctctccctgttccgaacctg	1499
OY	1455	ccggttaccggaataccggttcgcgcctttccctcttgagggaagcgtgtggcgtttctcataagc	1514
Db	1500	ccggttaccggaataccggttcgcgcctttccctcttgagggaagcgtgtggcgtttctcataagc	1559
OY	1515	tcacgcgttagtgatactccagttctggtgtatagtgcttctgcgtcccaagctgtggcgtgtgcaac	1574
Db	1560	tcacgcgttagtgatactccagttctggtgtatagtgcttctgcgtcccaagctgtggcgtgtgcaac	1619
OY	1575	gaaccccccgtlcaagccgcgcgcgtgcgcgttataccggttaactatcgtctctgtagctcaac	1634
Db	1620	gaaccccccgtlcaagccgcgcgcgtgcgcgttataccggttaactatcgtctctgtagctcaac	1679
OY	1635	ccggttaagaacagacttatctgcacactgtgcgcgcgcgcacactgtgtataacagattatgacgaagc	1694
Db	1680	ccggttaagaacagacttatctgcacactgtgcgcgcgcgcacactgtgtataacagattatgacgaagc	1739

OY	1695	agcatatgaacggtgtgtacaaaggtcttcgttaagttgtgtgccaactacgcgtcaacataga	1754
Db	1740	aggtatctaaagcggtgtgtacaaagattctcttgaagtgtgtgtgccaactacgcgtcaacataga	1799
OY	1755	aggacgaattattgttactctgcgtctgtcgttgaagccagttactcttcggaaaaagattgt	1814
Db	1800	aggacgaattattgttactctgcgtctgcgtcgtgaagccagttactcttcggaaaaagattgt	1859
OY	1815	agccttgcattccgcgaacaacaacacacgcgtgtgaagcggtgttttttgcattgbcgaacag	1874
Db	1860	agccttgcattccgcgaacaacaacacacgcgtgtgaagcggtgttttttgcattgbcgaacag	1919
OY	1875	caagattacgcgcgaagaaaaaagagatctcaagaagatcccttgcattcttctacggggtct	1934
Db	1920	caagattacgcgcgaagaaaaaagagatctcaagaagatcccttgcattcttctacggggtct	1979
OY	1935	gaagctcaatgttggaacggaaaactcaagttaaaggaatttvtgtatgaagcggatatacttt	1994
Db	1980	gaagctcaatgttggaacggaaaactcaagttaaaggaatttvtgtatgaagatatacttaaaag	2039
OY	1995	gaatgatattaga 2007	
Db	2040	atcttaacactaga 2052	

RESULT	9
AAV38297	ID AAV38297 standard; DNA; 5446 BP.
XX	
XX	
AC	AAV38297;
XX	
DT	26-Oct-1998 (first entry)
XX	
DE	Plasmid pCDNA3.
XX	
XX	Plasmid pCDNA3; pneumococcal surface protein A; PspA, infection;
KW	Streptococcus pneumoniae; sepsis; otitis media; meningitis;
KW	bacteraemia; pneumonia; vaccine; genetic immunisation; ss.
XX	
OS	Chimeric - human cytomegalovirus.
XX	
PN	W09824927-A1.
PD	11-JUN-1998.
XX	
PF	04-DEC-1997; 97WO-US22847.
XX	
PR	04-DEC-1996; 96US-0759505.
XX	
PA	(UYAL-) UNITV ALABAMA.
XX	
PI	Briles DE, Curiel DT, McDaniel LS;
DR	WPI; 1998-333343/29.
XX	
PT	Plasmid containing pneumococcal epitope for expression in eukaryotic
PT	cells - useful for eliciting immunological response to pneumococcal
PT	infection or sepsis
XX	
PS	Example 1; Fig 1B1 to 1B-5; 47pp; English.
XX	
CC	This is the DNA sequence of plasmid pCDNA3 (Invitrogen). A portion
CC	of the gene (see AAV38298) that codes for respiratory syncytial virus
CC	glycoprotein G (RSVG) has been amplified, digested with KpnI and
CC	ligated into KpnI-digested pCDNA3 upstream of the multiple cloning
CC	site of pCDNA3 and downstream of the human cytomegalovirus
CC	immediate early (HCMV-IE) promoter to create pGT41. A full-length
CC	coding sequence of Streptococcus pneumoniae Rxi pneumococcal
CC	surface protein A (PspA) was then inserted into pGT41 to create
CC	a fusion between rsvg and PspA. Intramuscular immunisation of
CC	BALB/c mice with the resulting plasmid, designated PKSD2601,
CC	induced protection against an otherwise lethal challenge with a
CC	capsular type 3 pneumococcus. A plasmid for expression of

CC pneumococcal epitope DNA in eukaryotic cells is claimed. The
CC plasmid includes a promoter for driving expression in a eukaryotic
CC cell (e.g. HCMV-IE), DNA encoding a leader sequence (e.g. of RSVG)
CC which facilitates expression, translation through or transport of
CC the expression product in a eukaryotic cell membrane, and DNA
CC encoding a pneumococcal epitope such as PSPA. The invention also
CC provides a vaccine comprising the plasmid and a suitable carrier or
CC diluent, and optionally one or more cytokines or DNA encoding them,
CC or a bacterial delivery system. The vaccine is used to elicit an
CC immunological response in a host, including humans, susceptible to
CC pneumococcal infection or sepsis. The plasmid can also be used to
CC express a pneumococcal epitope of interest in vitro.

XX Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 other;

Query Match 13.3%; Score 1082; DB 19; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.5e-153;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatgtacggcgccagatatacgcgttgacattgattatgacagtattatgaatca 60
DB 210 gatgtacggcgccagatatacgcgttgacattgattatgacagtattatgaatca 269
QY 61 attacggggatcattagttacatagccatatalygaagttccggttacataactacgta 120
DB 270 attacggggatcattagttacatagccatatalygaagttccggttacataactacgta 329
QY 121 aatgcccgcctgctgtgacgcgcgaacgaccccgccattgacgttcaataatgacgtat 180
DB 330 aatgcccgcctgctgtgacgcgcgaacgaccccgccattgacgttcaataatgacgtat 389
QY 181 gtcccatagtaacgcgaataggaagcttccattgacgcgaatggtgtgacattacag 240
DB 330 gtcccatagtaacgcgaataggaagcttccattgacgcgaatggtgtgacattacag 449
QY 241 taaactgcccacttggcagttacatcaatgtaacatgtaacatgtaacatgtaac 300
DB 450 taaactgcccacttggcagttacatcaatgtaacatgtaacatgtaacatgtaac 509
QY 301 gtccatagtaacgcgaataggaagcttccattgacgcgaatggtgtgacattacag 360
DB 510 gtccatagtaacgcgaataggaagcttccattgacgcgaatggtgtgacattacag 569
QY 361 cctacttggcagttacatcaatgtaacatgtaacatgtaacatgtaacatgtaac 420
DB 570 cctacttggcagttacatcaatgtaacatgtaacatgtaacatgtaacatgtaac 629
QY 421 cagttacatcaatggtgagtagcggttggactcaacgggagttcccaagttccacccc 480
DB 630 cagttacatcaatggtgagtagcggttggactcaacgggagttcccaagttccacccc 689
QY 481 attgacgtcaatgggaggtttgtttggcaccaaatcaacgggagttcccaagttcc 540
DB 690 attgacgtcaatgggaggtttgtttggcaccaaatcaacgggagttcccaagttcc 749
QY 541 aacaaactcgcccatatggcgaatggcggtgagcggtgtaacggtggaggtcctata 600
DB 750 aacaaactcgcccatatggcgaatggcggtgagcggtgtaacggtggaggtcctata 809
QY 601 agcagagctctctgcttaactagagaaccactgcttaactgcttalcgaataatacag 660
DB 810 agcagagctctctgcttaactagagaaccactgcttaactgcttalcgaataatacag 869
QY 661 actactatagggagagccaagcttggtaacgagctcgatccactatgtaacggcgcca 720
DB 870 actactatagggagagccaagcttggtaacgagctcgatccactatgtaacggcgcca 929
QY 721 ggtgtgtgaattcttcagatatccatcacactggtcgccgctcgagacgtactagag 780
DB 930 ggtgtgtgaattcttcagatatccatcacactggtcgccgctcgagacgtactagag 989
QY 781 ggcctatctatagtgatcccaaatgtcagtagcgtcgtatcagcctcgtactgtgct 840

DB 990 ggcctatctatagtgatcccaaatgtcagtagcgtcgtatcagcctcgtactgtgct 1049
QY 841 tctagtgcagagccatctgtgtttgtccctcccccgcgttcccttgaacctggaaggt 900
DB 1050 tctagtgcagagccatctgtgtttgtccctcccccgcgttcccttgaacctggaaggt 1109
QY 901 gccactccacatgctcttctcctaataaagtgggaattgcatctgctgtgtgagtag 960
DB 1110 gccactccacatgctcttctcctaataaagtgggaattgcatctgctgtgtgagtag 1169
QY 961 tctcattctatcttggtgggtggtggtggtggtggtggtggtggtggtggtggtggt 1020
DB 1170 tctcattctatcttggtgggtggtggtggtggtggtggtggtggtggtggtggtggt 1229
QY 1021 aataagcagcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1080
DB 1230 aataagcagcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1289
QY 1081 tg 1082
DB 1290 tg 1291

RESULT 10
AA53871
ID AA53871 standard; DNA; 5245 BP.
XX
AC AA53871;
XX
DT 03-JAN-2001 (first entry)
XX
DE Expression vector pRIG22b.
XX
KW Vector; endogenous gene; activation; over-expression;
KW erythropoietin; growth hormone; drug discovery;
KW granulocyte colony stimulating factor; ds.
OS Synthetic.
XX
PN WO200049162-A2.
XX
PD 24-AUG-2000.
XX
PF 22-FEB-2000; 2000MO-US04429.
XX
PR 19-FEB-1999; 99US-0253022.
PR 08-MAR-1999; 99US-0263814.
PR 26-MAR-1999; 99US-0276820.
XX
PA (ATHE-) ATHERSYS INC.
XX
PI Harrington JJ, Sherf B, Rundlett S;
XX
DR WPI; 2000-549276/50.
XX
PT Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
XX
PS Disclosure; Fig 16; 240pp. English.
XX
XX New methods, vectors and cells are described for non-targeted
XX activation and over-expression of endogenous genes. The vector
XX constructs comprise transcriptional regulatory sequences (TRS) and
XX unpaired splice donor sequences (USDS), preferably the vectors
XX comprise (in sequential order) a TRS, an USDS, a rare cutting
XX restriction site (KCRS) and a linearization site (LS) with a second
XX TRS linked to a selectable marker (SM) lacking a polyadenylation
XX signal. The methods, vectors and cells comprising the vectors may
XX be used for the non-targeted activation and over-expression of
XX endogenous genes, e.g. for the production of proteins (including
XX erythropoietin, growth hormone or granulocyte-colony stimulating

Db 1987 gatcttcacctaga 2000

RESULT 11

AAV40007
ID AAV40007 standard; DNA: 4026 BP.

AAV40007:

15-FEB-1999 (first entry)

Plasmid pCTMI.

E2F; transcription factor; human; retinoblastoma protein RB;

bladder cancer; restenosis; angioplasty; diabetic retinopathy;

thyroid hyperplasia; Grave's disease; psoriasis;

benign prostatic hypertrophy; Li-Fraumeni syndrome;

peripheral vascular disease; therapy; plasmid pCTMI; ss.

Chimeric - cytomegalovirus.

Chimeric - mastadenovirus.

Chimeric - bacteriophage T7.

Chimeric - rhesus macaque polyoma virus.

Chimeric - Bos taurus.

Key Location/Qualifiers

promoter 209..864 /tag= a

misc-feature 907..1074 /note= "CMV promoter"

Intron 1075..1253 /tag= b

promoter 1305..1322 /note= "hybrid SV40 late intron"

misc-feature 1851..4026 /note= "SP6 promoter"

CDS /tag= e

/note= "pUC19 backbone H3 to AatII"

/tag= i

/note= "AMP-ORF"

MO9821228-A1.

22-MAY-1998.

13-NOV-1997; 97MO-US21821.

14-FEB-1997; 97US-0801092.

15-NOV-1996; 96US-0751517.

(CANJ-) CANJ1 INC.

Antelman D, Gregory RJ, Wills KN;

WPI; 1998-297858/26.

New fusion polypeptide of, e.g. transcription factor - used to

treat, e.g. hyper-proliferative disease such as cancer and

restenosis

Example 1; Fig 6; 91pp; English.

This is the nucleotide sequence of pCTMI, a plasmid that was

constructed from pCTM (see AAV40006) by digesting pCTM with XhoI and

CC retinoblastoma protein (BP, see AAV62465) and E2F transcription
CC factor (see AAV62464). Such fusion proteins, particularly expressed
CC from gene therapy vectors, are used to treat hyperproliferative
CC conditions, specifically cancer (particularly of the bladder) or
CC restenosis. They are more effective in repressing transcription of
CC the E2F promoter than RB alone and cause cell-cycle arrest in a
CC variety of cells.
XX
SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 other;

Query Match 13.2%; Score 1070.2; DB 19; Length 4026;
Best Local Similarity 69.7%; Pred. No. 1.5e-151;
Matches 1919; Conservative 0; Mismatches 88; Indels 746; Gaps 5;

QY 1 gatgacggcgacgatatagcagctgtagcattgattgactgattatgaatgaac 60
Db 210 gatgacggcgacgatatagcagctgtagcattgattgactgattatgaatgaac 269
QY 61 attacggggtcattagttcattagcccatatagagttcgcggttaactaactagga 120
Db 270 attacggggtcattagttcattagcccatatagagttcgcggttaactaactagga 329
QY 121 aatggcccgctgctgtagccgccaacgaccccgccattgagctgaataatagcat 180
Db 330 aatggcccgctgctgtagccgccaacgaccccgccattgagctgaataatagcat 389
QY 181 gtcccatagtaacgacgacataggagcttccattgacgtcaatggttgaattatgacg 240
Db 390 gtcccatagtaacgacgacataggagcttccattgacgtcaatggttgaattatgacg 449
QY 241 taaactgccacttggcagatcatcatcaagtatcatatgccaagtacgccccattgac 300
Db 450 taaactgccacttggcagatcatcatcaagtatcatatgccaagtacgccccattgac 509
QY 301 gtcaatgacggttaatggccgctgcatattgccaagtacatgacttggagactt 360
Db 510 gtcaatgacggttaatggccgctgcatattgccaagtacatgacttggagactt 569
QY 361 cctacttggcagatcatcatgattatgacatgcatatgacatgattgattgacggtt 420
Db 570 cctacttggcagatcatcatgattatgacatgcatatgacatgattgattgacggtt 629
QY 421 cagttacatcaatggcggtgtagacggtttgactcagcggtgattccaagttccacccc 480
Db 630 cagttacatcaatggcggtgtagacggtttgactcagcggtgattccaagttccacccc 689
QY 481 attgacgtcaatggaggtttgtttggcaccacaaatcaacgggacttccaaatgctg 540
Db 690 attgacgtcaatggaggtttgtttggcaccacaaatcaacgggacttccaaatgctg 749
QY 541 aacaactccgcccattttagcgaataggcggttagcggtgacggttggaggtctatata 600
Db 750 aacaactccgcccattttagcgaataggcggttagcggttggaggtctatata 809
QY 601 agcagagctctctggttaactagagaacccaactgcttactggtctatagaattatagc 660
Db 810 agcagagctctctggttaactagagaacccaactgcttactggtctatagaattatagc 869
QY 661 actcaactataggagagaccgaactgg----- 687
Db 870 actcaactataggagagaccgaactggcggtgtaccacactctctccgcatgctgctg 929
QY 688 ----- 687
Db 930 cgaaggccagctgttggtcgcggttgaggaacaaactctccggtcttccagtaactc 989
QY 688 ----- 687
Db 990 ttgagtcggaacccgctgcctccgaacggttactcgcacacggggagacctgagcagt 1049
QY 688 ----- 687

CC of the virus. The multivalent vaccine protects against pathologies
CC resulting from exposure to one or several hantaviruses. The
CC advantages of the composition are that the risks of pathogenic
CC infection are not present as live vaccines are not used and
CC production of the composition does not require growth or use of
CC hantavirus which can be spread by aerosol transmission. To
CC determine the efficacy of the M gene segment of Seoul virus to
CC provide protective immunity the gene segment was subcloned into the
CC naked DNA expression vector pMRG7077 downstream of the
CC cytomegalovirus immediate early promoter to yield pMRG/SEO-M. The
CC construct was then used to challenge 6 to 8 week old female BALB/c
CC mice. Mice vaccinated with pMRG/SEO-M produced neutralising
CC antibodies.

SQ Sequence 8002 BP; 2208 A; 1788 C; 1818 G; 2188 T; 0 other;

Query Match 13.0%; Score 1058.4; DB 21; Length 8002;
Best Local Similarity 94.4%; Pred. No. 7.8e-150;
Matches 1121; Conservative 0; Mismatches 36; Indels 31; Gaps 1;

QY 820 tgatcagcctgactgctgcttctagttgacagcactctgttggccctccctg 879
DB 6706 tgatcagcctgactgctgcttctagttgacagcactctgttggccctccctg 6765
QY 880 ccttcctgacccctggaagtgccacccctccactgtccttccataaataagagaatt 939
DB 6766 ccttcctgacccctggaagtgccacccctccactgtccttccataaataagagaatt 6825
QY 940 gcatcgactgtctgtagttagtgcattcttctgagggtggtggtggtggaagcagc 999
DB 6826 gcatcgactgtctgtagttagtgcattcttctgagggtggtggtggtggaagcagc 6885
QY 1000 aaggaggagaggttggaagacaatagcagcagctggtggtggtggtggtggtggt 1059
DB 6886 aaggaggagaggttggaagacaatagcagcagctggtggtggtggtggtggtggt 6945
QY 1060 tctgagcgagaaagacacagctgcatatgatcgcaacgcggtggaagcggtt 1119
DB 6946 tctgagcgagaaagacacagctgcatatgatcgcaacgcggtggaagcggtt 6974
QY 1120 tgcgtattggtgctctcctcctcctcctcctcctcctcctcctcctcctcctc 1179
DB 6975 gacagctgacttagaattgtctcctcctcctcctcctcctcctcctcctcctc 7034
QY 1180 tgcgtgagcggtatcagctcactcaaaagcggtatcaggtttaccacaatagagg 1239
DB 7035 tgcgtgagcggtatcagctcactcaaaagcggtatcaggtttaccacaatagagg 7094
QY 1240 ataacgcaggaagaacatgtgagcaaaagcgcaagaagcgcaagcgtataaagg 1299
DB 7095 ataacgcaggaagaacatgtgagcaaaagcgcaagaagcgcaagcgtataaagg 7154
QY 1300 ccgcgtgctggtcttcttcataagctcgccctcctcctcctcctcctcctcctc 1359
DB 7155 ccgcgtgctggtcttcttcataagctcgccctcctcctcctcctcctcctcctc 7214
QY 1360 gctcaagtcagagtggtggaaccccgacagacataaagataccagcggttccctc 1419
DB 7215 gctcaagtcagagtggtggaaccccgacagacataaagataccagcggttccctc 7274
QY 1420 gaaactcctcgtgctcctccttcccgacctcctcctcctcctcctcctcctc 1479
DB 7275 gaaactcctcgtgctcctccttcccgacctcctcctcctcctcctcctcctc 7334
QY 1480 tctcctcctcgggaagcgtggtccttctcctaagtctcagcgtgtatctcagttcg 1539
DB 7335 tctcctcctcgggaagcgtggtccttctcctaagtctcagcgtgtatctcagttcg 7394
QY 1540 tgtatgctcgtcgtcctcagagctggtgtgtgcaagaaccccggttaagccgagcgt 1599
DB 7395 tgtatgctcgtcgtcctcagagctggtgtgtgcaagaaccccggttaagccgagcgt 7454

QY 1600 ggccttaccgtaactacgtcttgtagtccaacccggtgaagacagactatcgccac 1659
DB 7455 ggccttaccgtaactacgtcttgtagtccaacccggtgaagacagactatcgccac 7514
QY 1660 tggcagcagcactgtgtacagagatcagcagcgaggtatgtagcggtgtctaagagt 1719
DB 7515 tggcagcagcactgtgtacagagatcagcagcgaggtatgtagcggtgtctaagagt 7574
QY 1720 tcttgaagtgtggtcctactacagcctacacataagagacagatttgatctgcgtc 1779
DB 7575 tcttgaagtgtggtcctactacagcctacacataagagacagatttgatctgcgtc 7634
QY 1780 tgcgtgaagccagttacccctcgtgaaagagtggtgtagctctgtacccggcaacaaca 1839
DB 7635 tgcgtgaagccagttacccctcgtgaaagagtggtgtagctctgtacccggcaacaaca 7694
QY 1840 ccgctgtgtagcggtgtgttttttttttgcgaagcacaattacgcgcgaagaaagagat 1899
DB 7695 ccgctgtgtagcggtgtgttttttttttgcgaagcacaattacgcgcgaagaaagagat 7754
QY 1900 ctcaagaagatcccttgatcttctcaggggtctgacgctcagtggaacgaaactcac 1959
DB 7755 ctcaagaagatcccttgatcttctcaggggtctgacgctcagtggaacgaaactcac 7814
QY 1960 gttaaaggatttggctcaltgagcggatataatttgaatgtattaga 2007
DB 7815 gttaaaggatttggctcaltgagatatacaaaagagattcaactaga 7862

RESULT 15
AAA27831
ID AAA27831 standard; DNA; 3796 BP.

XX AC AAA27831;

XX DT 12-SEP-2000 (first entry)

XX DE Vector plasmid pCMV-MC1.

XX KW North American PRRS virus; Nidovirales virus; pig; swine; vaccine;
KW pCMV-S-P129; ss.

XX OS Chimeric - Porcine reproductive and respiratory syndrome virus.

XX OS Chimeric - Human cytomegalovirus.

XX EN EPI018557-A2.

XX PD 12-JUL-2000.

XX PF 25-NOV-1999; 99EP-0309409.

XX PR 22-DEC-1998; 98US-0113345.

XX PA (PRIZ) PRIZER PROD INC.

XX PI Calvert JG, Welch SW, Shepard MG;

XX DR WPI: 2000-444364/39.

XX PT New polynucleotide encoding an infectious RNA molecule of a North
PT American porcine reproductive and respiratory syndrome virus for use as
PT a vaccine in protecting swine and other animals from infection by a
PT pathogen -

XX PS Example 4; Page 43-44; 53pp; English.

XX CC The present sequence is that of eukaryotic expression vector
CC plasmid pCMV-MC1. The plasmid is derived from pCMVbeta by
CC replacing the lacZ coding sequence with a linker containing
CC multiple restriction sites. Modification of the human
CC cytomegalovirus (CMV) immediate early promoter was accomplished by
CC substituting the sequence between SacI and the 2nd NotI site of
CC pCMV-MC1 with a synthetic linker. North American porcine

Qy 1681 ggattagcagagcgaggtatgtatgagcggtgctacagagttcttgaagtgtgacctact 1740
 |||
 Db 1947 ggattagcagagcgaggtatgtatgagcggtgctacagagttcttgaagtgtgacctact 2006
 Qy 1741 acggtacactagaagagcaatatttggatctggtctgtgtaagccagttaacctcg 1800
 |||
 Db 2007 acggtacactagaagagcaatatttggatctggtctgtgtaagccagttaacctcg 2066
 Qy 1801 gaaaaaggtgtgtagctcttgatccgcaacaaccaccgctgtgtagcggtgtttt 1860
 |||
 Db 2067 gaaaaaggtgtgtagctcttgatccgcaacaaccaccgctgtgtagcggtgtttt 2126
 Qy 1861 ttgtttgcaagcagcagattacgcgcagaaaaaagga tctcaagaagatccttggatct 1920
 |||
 Db 2127 ttgtttgcaagcagcagattacgcgcagaaaaaagga tctcaagaagatccttggatct 2186
 Qy 1921 ttctacgggtctgagcgtcagtggaagaaaaactcaagtttaaggattttgtcatga 1980
 |||
 Db 2187 ttctacgggtctgagcgtcagtggaagaaaaactcaagtttaaggattttgtcatga 2246
 Qy 1981 gcggtacacata ttgaa tttataga 2007
 |||
 Db 2247 gattatcaaaaaggtatcttcaactaga 2273

Search completed: November 3, 2001, 20:43:59
 Job time: 16975 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:49:24 : Search time 154.79 Seconds
(without alignments)
9924.800 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatgtacgggcagatatatcac.....gcgtgtcgtacgtacgcagcgtc 8115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCRTUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240.4	15.3	3853	3 US-08-801-092-5	Sequence 5, App11
2	1068.6	13.2	4026	3 US-08-801-092-19	Sequence 19, App1
3	1053.6	13.0	4326	4 US-08-760-615-7	Sequence 7, App11
4	981.8	12.1	6253	3 US-08-893-327-15	Sequence 15, App1
5	981.8	12.1	6280	3 US-08-893-327-17	Sequence 17, App1
6	981.8	12.1	6280	3 US-08-893-327-19	Sequence 19, App1
7	973	12.0	4283	1 US-08-343-401A-3	Sequence 3, App11
8	973	12.0	4283	1 US-08-445-265A-1	Sequence 1, App11
9	973	12.0	4283	3 US-08-990-442-1	Sequence 1, App11
10	921.8	11.4	6285	1 US-08-467-420A-49	Sequence 49, App1
11	921.8	11.4	6285	1 US-08-470-110A-49	Sequence 49, App1
12	921.8	11.4	6285	1 US-08-667-769A-49	Sequence 49, App1
13	921.8	11.4	6285	2 US-08-940-371-49	Sequence 49, App1
14	921.8	11.4	6285	3 US-08-637-647-49	Sequence 49, App1
15	921.8	11.4	6285	5 PCF-US95-17082A-49	Sequence 49, App1
16	906.4	11.2	3130	4 US-09-038-141-1	Sequence 1, App11
17	906.4	11.2	3789	4 US-09-075-019-8	Sequence 8, App11
18	906.4	11.2	3918	2 US-08-495-500-2	Sequence 2, App11
19	906.4	11.2	4249	1 US-08-801-092-33	Sequence 33, App1
20	906.4	11.2	4453	1 US-08-770-761A-4	Sequence 4, App11
21	906.4	11.2	4454	2 US-08-738-172-2	Sequence 2, App11
22	906.4	11.2	4525	2 US-08-613-861-2	Sequence 6, App11
23	906.4	11.2	4540	1 US-08-770-761A-6	Sequence 6, App11
24	906.4	11.2	4824	2 US-08-485-139-5	Sequence 5, App11
25	906.4	11.2	4824	3 US-08-750-357-5	Sequence 5, App11
26	906.4	11.2	4883	1 US-08-064-121-4	Sequence 4, App11
27	906.4	11.2	4883	1 US-08-318-772A-1	Sequence 1, App11

28	906.4	11.2	4883	1 US-08-478-015-4	Sequence 4, App11
29	906.4	11.2	4883	3 US-08-475-975-4	Sequence 4, App11
30	906.4	11.2	4883	3 US-09-084-889-4	Sequence 4, App11
31	906.4	11.2	4949	3 US-09-138-024-22	Sequence 22, App11
32	906.4	11.2	5399	1 US-08-064-121-1	Sequence 1, App11
33	906.4	11.2	5399	1 US-08-478-015-1	Sequence 1, App11
34	906.4	11.2	5399	3 US-08-475-975-1	Sequence 1, App11
35	906.4	11.2	5399	3 US-09-084-889-1	Sequence 1, App11
36	906.4	11.2	5618	3 US-08-799-569-1	Sequence 1, App11
37	906.4	11.2	5620	1 US-08-104-072B-9	Sequence 9, App11
38	906.4	11.2	5620	1 US-08-351-413-3	Sequence 3, App11
39	906.4	11.2	5620	2 US-09-025-583-3	Sequence 3, App11
40	906.4	11.2	5642	1 US-08-318-772A-2	Sequence 2, App11
41	906.4	11.2	5653	1 US-08-073-836-3	Sequence 3, App11
42	906.4	11.2	5653	1 US-08-233-277-1	Sequence 1, App11
43	906.4	11.2	5703	1 US-08-467-420A-50	Sequence 50, App1
44	906.4	11.2	5703	1 US-08-470-110A-50	Sequence 50, App1
45	906.4	11.2	5703	1 US-08-667-769A-50	Sequence 50, App1

ALIGNMENTS

RESULT 1
US-08-801-092-5
; Sequence 5, Application US/08801092
; Patent No. 6074850
; GENERAL INFORMATION:
; APPLICANT: Antelman, Douglas
; APPLICANT: Gregory, Richard J.
; APPLICANT: Mills, Kenneth N.
; TITLE OF INVENTION: Tissue Specific Expression of
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,092
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,517
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 703-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3853 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 209..250
; FEATURE:
; NAME/KEY: CDS

LOCATION: 254..289
FEATURE:
NAME/KEY: CDS
LOCATION: 293..505
FEATURE:
NAME/KEY: CDS
LOCATION: 509..514
FEATURE:
NAME/KEY: CDS
LOCATION: 518..520
FEATURE:
NAME/KEY: CDS
LOCATION: 524..658
FEATURE:
NAME/KEY: CDS
LOCATION: 662..691
FEATURE:
NAME/KEY: CDS
LOCATION: 695..748
FEATURE:
NAME/KEY: CDS
LOCATION: 752..781
FEATURE:
NAME/KEY: CDS
LOCATION: 785..829
FEATURE:
NAME/KEY: CDS
LOCATION: 1132..1134
FEATURE:
NAME/KEY: CDS
LOCATION: 1138..1149
FEATURE:
NAME/KEY: CDS
LOCATION: 833..862
US-08-801-092-5

Query Match 15.3%; Score 1240.4; DB 3; Length 3853;
Best Local Similarity 74.1%; Pred. No. 1e-268;
Matches 1911; Conservative 0; Mismatches 96; Indels 573; Gaps 4;

QY 1 gatgacggcgccagatacgcgtgacatgatatgactagtagtataatgaatca 60
DB 210 GATGACGGCGCCAGATATACGGCTGACATGATTGACTAGTTATTAATGAATCA 269
QY 61 attaggggtattatgctatagcccatatagaggtccgggttaacttaactcgta 120
DB 270 ATTAGGGGTCTATTAGTTCATAGCCCATATATGAGTTCCGGTACATACCTACGGTA 329
QY 121 aatggcccgctgctgacgcgcacacgaaccccgcccatgacgcaataatgacgat 180
DB 330 AATGGCCGCGCTGCTGACCGCCCAACGACCCCGCCCATGACGTCATATGACGTA 389
QY 181 gtcccatagtaacgcacaaataggaacttcaattgaacgtcaatgggtgactacttaacgg 240
DB 390 GTTCCCATAGTAACGCCCAATAGGGACTTTCATGTGACGTCAATGGGTGACTATTACGG 449
QY 241 taaatgcccacttgagcgtatcaatagtgatacgaagtaagccccccttaac 300
DB 450 TAAATGCCCACTTGGCGATACATCAAGTATCATATGCAAGTACGCCCTTATGAC 509
QY 301 gtcaatgacgtlaaatgcccgcctgcatltagccagtaacatlaacctlaaggactt 360
DB 510 GTCAATGACGTAATATGGCCCGCTGGCATTTATGCCCAATACCTTATGGGACTTT 569
QY 361 cctacttgacgtatcatctagctatagcgtatcatctaccatgagtgagcggtttgg 420
DB 570 CCTACTTGGCAGTACATCATCTATTAGTCATCGCTATTACCATGATGGGTTTGG 629
QY 421 cagtcacatcaatggcggtgagtagcggtttgactcaacggggatcccaagttccacccc 480
DB 630 CAGTCAATCAATGGCGGTGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC 689

QY 481 attgacgtcaatlggaggttctgtttgacacaaaatcaacgggacttccaaaatgtcgt 540
DB 690 ATTGACGTCAATGGGAGTTGTTTGGCACCAAAATCAACGGGACTTTCAAAATGTGTGT 749
QY 541 aacaactcgcggccattgacgcaaatlggcggttaggcgtgtaaggttggaggtcatala 600
DB 750 AACAACTCGGCCCTTATGACGCAAAATGGGCGGTACGCTGTAGGTTGGAGGTCTATATA 809
QY 601 acaagactctctgcttaactagagaaccacactgcttaactgctatagcaatlaacg 660
DB 810 AGCAGAGCTCTCTGGCTTAAGTACAGAACCCACTGCTTACTGGCTTATGGAATTAATACG 869
QY 661 actcactaaggagagaccacagctt----- 685
DB 870 ACTCACTATAGGAGAACCCAGAGCTTCGGCGGATACACTCTCTTCCCATGCTGTCTG 929
QY 686 ----- 685
DB 930 CGAGGCCAGCTGTGCGCTCGCGGTTGAGAGCAAACTTTCGCGCTTCCAGTACTC 989
QY 686 -----ggtaccgagctcgagatccactagtaacgagcgccagtgatgc 726
DB 990 TTGCAATCGGAACCCGCTGCGCTCCGAACGGTACTCCGCCACCGAGGACCTGACGAGT 1049
QY 727 tggaaattctgagatatacatcaacactgagcgccgc-----tcgag 767
DB 1050 CCGCATCGACGGGATCGGAAACCTCTCGAGCGCGCGCTGCACTAGACGAATTCGCG 1109
QY 768 catgacatagaagggccctatctatagtgatcaactaaatgcttagagctgctgaacgc 827
DB 1110 TACGATATCGATGGGCGCTTATGCTATGTCACCTAAATGCTAGAGCTCGCTGATCAGC 1169
QY 828 ctcgactgtgcctctagcttagctgacacatcgtgtgttgcgcctcccgatgcctcct 887
DB 1170 CTCGACTGTGCTTCTTACTGTGCGACCACTGTGTGTTTGCCCTCCCGGCTTCCTT 1229
QY 888 gacccttgaaggtgacactccactgctccttccctaataatgaggaatgacatgcga 947
DB 1230 GACCTGGAAGGTGCGACCTCCACTGCTCTTCTTAATAATGAGAAATGTCATCGCA 1289
QY 948 ttgtctgagtaggtgtcaattctatctcgtgggggttgggttgggacgaagaaggggga 1007
DB 1290 TTGCTGAGTAGTGTCTCATTTCTATTTCTGGGGGTGGGGTGGGCGACGACGAAGGGGA 1349
QY 1008 gatttgggaagaacaataagcaagcagtcgcygga----- 1040
DB 1350 GATTGGGAAGCAATAGCCGAATGACCGACGACGACGCCCAACCTGCCATCAGCAG 1409
QY 1041 ----- 1040
DB 1410 ATTGCAATTCACCGCGCTTCTATGAAAGGTTGGGCTCGGAATGTTTTCCGGAGC 1469
QY 1041 ----- 1040
DB 1470 CCGGCTGATGATCTCCAGCGCGGGGATCTCATGCTGAGTTCTTCGCCCACT 1529
QY 1041 ----- 1040
DB 1530 TGTTATTGACGCTTATAATGTTTACAAATAAAGCAATAGCATCACAAAATTCACAATA 1589
QY 1041 ----- 1040
DB 1590 AAGCAATTTTTCACCTGCACTTCTAGTTGTGTTTGTCCAAACTCATCATATGATCTATC 1649
QY 1041 ----- 1040
DB 1650 ATGTCTGATACCGTGAACCTCTAGCTAGAGCTTGGCGTATCATGATAGCTGTTTC 1709
QY 1041 ----- 1040
DB 1710 CTGTGTGAATTTGTTATCGCTCACAATTTCCACACAAATACGAGCCGGAAGCATAAAGT 1769
QY 1041 -----tgcggtgggtctatag 1057

Db 1770 GTAAGCCTGGGCTGCTATAGTAGCTAACTACATTAAATTGGGTGGCTCACTGC 1829
Qy 1058 ctctc-----gagcggaagaaccagctgcatlaatgaatcgccaaagcgcg 1107
Db 1830 CCGCTTTCAGTCCGAGAACCTCTCGTCAGCTGATTAATGAATGGCCAGCGCGG 1889
Qy 1108 gtagagagcggttggtggtatgtggcgtcttcctccgtcccaatgactgctgct 1167
Db 1890 GGAAGAGCGGTTGGCTATTGGGCGCTTCCTTCCTCGCTACATGACTCCCTCGCT 1949
Qy 1168 cggctgctcgagctgagcgagcgtlaacgtcaaccgaagcggtcaatcaggttcca 1227
Db 1950 CGGTCTGTGGCGTGGCGCCAGCGGTATCACTCACTCAAAAGCGGTATATCGGTATCCA 2009
Qy 1228 cagaatcagggagtaacgaagaagaacatgtgacaaaggccagcaaaagcgcaaga 1287
Db 2010 CAGAATCAGGGGATACGACGAGAAACATGTGACCAAAAGGCCAGCAAAAGGCCAGA 2069
Qy 1288 accgtataaaggcggtgtgctgtggttttccatagctccgccccctgaagagatc 1347
Db 2070 ACCGTAAAAAGCGCGTGTGCTGGCTTTTCCATAGCTCCGCCCTGACGAGCATC 2129
Qy 1348 acaaaatcagcgtcgaatcagaagtgagcgaaccggaagactataaagatacag 1407
Db 2130 ACAAAATCGACCGCTCAAGTCAAGGTGGCGAAACCCGACAGCATATTAAGTATCCAGG 2189
Qy 1408 cgltcccccgtgaagctccctcgtagcgtctcgtltccgaccctcgcgctaccggt 1467
Db 2190 CGTTTCCCCGTGAAACCTCCCTCGCTCGCTCTCTCGACCCGCGCTTACCGGAT 2249
Qy 1468 acctgcgccttctcccttcgggaagcgttgagcgcttctcaatgctcaagctgtagt 1527
Db 2250 ACTGTGCCCTTTCCTCCCTCGGAAAGCGTGGCGCTTCTCATCTCACGCTTAGGT 2309
Qy 1528 atctcagtcggtgtagtgcgttcgtcccaagctggcgctgtagcgaagcccccgcttc 1587
Db 2310 ATCTCAGTTCCGGGTAGGTGCTTCGCTCCAGGTGGGCTGTGTGCAAGAACCCCGCTTC 2369
Qy 1588 agcccgacgcgtgagccttatccgtaactatcgtctgagtcacaaaccgtagaagac 1647
Db 2370 AGCCCGACCGCTGCGCTTATCCGCTGATACGTCTTGATGCTCAACCCGCTAAGACAG 2429
Qy 1648 acttaccgcacgtgagcgaagcgcactggttaacaggttagcgaagcggtatgtagcg 1707
Db 2430 ACTTATCGCACCTGCGACGACCATGTGTAACGATTAACAGAGCGATATGTAGGCG 2489
Qy 1708 gtgctacaagttcttgaagtgtagcctaactaagcgtactaagtagaagacagtttg 1767
Db 2490 GTGCTACAGATTCTTGAAGTGTGTGCTTAACGCTTAACGCTTAAGAGACGTAATTGG 2549
Qy 1768 gtatctgagctcgtcgtgaagccagttaccttcgaaaaagagttgtagctcttgacc 1827
Db 2550 GTATCTGCGCTGCTGTAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTTGTATCCG 2609
Qy 1828 gcaaaacaaaccccgctgtagcggtgtgtttttgttttgcaagcgacagattacgcga 1887
Db 2610 GCAAAACAAACCAACCGCTGATGAGGGGTGTTTGTGTTGCAAGCAGCATTAACCGCA 2669
Qy 1888 gaaaaaagaatcgaagaagatccttgatcttctctacgaggtctgacgcctcagttga 1947
Db 2670 GAAAAAAGAGATCTCAAGAGATCTTGTATCTTTCTACGGGGGTGTGACGCTCAATGGA 2729
Qy 1948 acgaaaaactcagcgtlaagggatttgatcatgagcgaatacatattgaatgattaga 2007
Db 2730 ACGAAAACTCAGGTAAAGGATTTTGTGATGATTAATCAAAAGAGATCTTACCTAGAG 2789

RESULT 2
US-08-801-092-19
; Sequence 19, Application US/08801092
; Patent No. 6074850
; GENERAL INFORMATION:

APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
TITLE OF INVENTION: Tissue Specific Expression of
TITLE OF INVENTION: Retinoblastoma Protein
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
FAX: 703-576-0300
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 4026 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 209..250
FEATURE:
NAME/KEY: CDS
LOCATION: 254..289
FEATURE:
NAME/KEY: CDS
LOCATION: 293..505
FEATURE:
NAME/KEY: CDS
LOCATION: 509..514
FEATURE:
NAME/KEY: CDS
LOCATION: 518..520
FEATURE:
NAME/KEY: CDS
LOCATION: 524..658
FEATURE:
NAME/KEY: CDS
LOCATION: 662..691
FEATURE:
NAME/KEY: CDS
LOCATION: 695..748
FEATURE:
NAME/KEY: CDS
LOCATION: 752..781
FEATURE:
NAME/KEY: CDS
LOCATION: 785..829
FEATURE:
NAME/KEY: CDS
LOCATION: 833..862

FEATURE: CDS
NAME/KEY: 1305..1307
LOCATION: 1305..1307
FEATURE: CDS
NAME/KEY: 1311..1322
LOCATION: 1311..1322
US-08-801-092-19

Query Match 13.2%; Score 1068.6; DB 3; Length 4026;
Best Local Similarity 69.7%; Pred. No. 3,2e-230;
Matches 1918; Conservative 0; Mismatches 89; Indels 746; Gaps 5;

QY 1 gatgtacggccagatatatacagcgttgaacatgatattgactagttatataatgaatca 60
DB 210 GATGTACGGCCAGATATACCGCTTGACATGATTGACTAGTATTATAGTAATCA 269
QY 61 attacgggtcattagttacatagaaccataatgaattccgcgttaacataactagta 120
DB 270 ATTACGGGTCATTAGTTCATATAGCCCATATATAGATTCCCGTTACATTAACGTA 329
QY 121 aatggccgcctgtgtgacgcgcacacacccgcgcacacacacacacacacacacac 180
DB 330 AATGGCCCGCTGTGTGACGCCACACACCCCGCCCATTTGACGTCAATATGACGTAT 389
QY 181 gtcccatagtaacagcaatagggactttccattgacgtcaatgggtgactattacg 240
DB 390 GTCCCATAGTAACGCCCAATAGGACTTTCATTGACGTCAATGGGTGACTATTTACGG 449
QY 241 taactgcgcacactgtgacgtatcaatgatatataatgaacgaagacacacacacac 300
DB 450 TAACTGCCACTGTGGCAGTACATCAAGTATATATATATATATATATATATATATAT 509
QY 301 gtcaatgacggttaaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
DB 510 GTCAATGACGGTAAATGCGCCGCGCATATATGCCGATACATACCTTATGGACCTT 569
QY 361 cctactgtgacgtatcaatcattagttacatcgctattacacgtgacgtgacgtgac 420
DB 570 CCTACTGTGACGTACATCTAGTATTATGCTATACGCTATTTACATGATGGGTTTGG 629
QY 421 cagtaacatcaatgagcggtgagtaggggttgaactacagggatcccaagtcacaccc 480
DB 630 CAGTACATCAATGGCGGTGATAGCGGTTTACACGCGGATTTCCAAAGTCTCCACCC 689
QY 481 attacgtcaatgaggtgtgtttttgtgacacaaatcaacagggacttccaaatgtcgt 540
DB 690 ATTGACGTCAATGGGAGTTGTTTGTGGACCAAAATCAACGGGACTTTCAAAATGTCT 749
QY 541 aacaactcgcgcacatgacgcaaatgagcggtgaggtgacgtgacgtgaggtgacgt 600
DB 750 AACCACTCCGCCCATTCACCAAAATGGCGGTAGCGCTGTAACGGTGTATATATA 809
QY 601 agcagagtcctgtgtacatagaagaacacacacacacacacacacacacacacacac 660
DB 810 AGCAGAGTCCTGTGGCTACTAGAGAACCCACTGCTTACTGGCTTATGAAATTAATACG 869
QY 661 actactactaagaggaacccaagcttg----- 687
DB 870 ACTGACTTATAGGAGAACCAAGCTTCGCGGGGTACCACTCTTCCGCAATCGCTGTCTG 929
QY 688 ----- 687
DB 930 CGAGGGCCAGCTGTGGGCTCGCGGTGAGACAAACTTTCGCGGTTCCTCCAGTACTC 989
QY 688 ----- 687
DB 990 TTGATGCGAAACCCGTCGCCCTCCGAACGCTACTCGCCACCGAGGAGCCTGACGAGT 1049
QY 688 ----- 687
DB 1050 CCGCATGACCGGATCGGAAACCTCTCGAGGAGACTGAAAAACAGAAAGTTAACTGGTA 1109

QY 688 -----taacgagctcggatccactagtaacgcccagtg 723
DB 1110 AGTTAGCTTTTGTCTTTTATTTATTTAGGTCGCCGATCCGTTGGTGCATAATCAAG 1169
QY 724 tgcgtgaattctgcagatat----- 743
DB 1170 AACTGCTCCTCAATGATGATGTCCTTACTTCTAGGCTGTACGAACTGTTACTTCTGC 1229
QY 744 -----ccatcacactggcgcgcgc-----tcgacatgcat 774
DB 1230 TCTAAAGCTGGGAAATGTATACCGCGCGCGGTGACGTCFAGACGAATTCGCTGATTA 1289
QY 775 ctagaagccctatctatagtgtaacctaaatgctaagctcgtgatacagctcgtact 834
DB 1290 TCGATGCGCCCTATCTATGATGTCACCTAAATGCTATAGCTCGCTGATCGCTCGACT 1349
QY 835 gtgccttagttgcaacacatctgtgtgttgccctccgcgcgcgcgcgcgcgcgcgcgc 894
DB 1350 GTGCCTTAGTTGCAACCACTGTGTGTGTGCCCCCTCCCGTCCCTTGACCCCTG 1409
QY 895 gaagtgccactcccaactgctccttcccaataaataaggaatgcatcgcattgctg 954
DB 1410 GAAGTGGCACTCCACGTCCTTCTTAATTAATGAGAAATGCAATGCAATGTCTG 1469
QY 955 agtaggtcattctatctcgtgggggtgggtggtgggacagcaaggggaggaattgg 1014
DB 1470 AGTAGGTCTCATTTCTATTTCTGGGGGTGGGTGGGCGACAGCAAGGAGAGATTTGG 1529
QY 1015 gaagacaaatagcagcagtcgctgggga----- 1040
DB 1530 GAAACAAATAGCCGAATGACCGACCAAGCGCCCAACTGCCATCAGAGATTTCGA 1589
QY 1041 ----- 1040
DB 1590 TTCCACCGCGCTTCTATGAAGGTTGGCTCGAATCGTTTCCGGAGCGCGCTG 1649
QY 1041 ----- 1040
DB 1650 GATGATTCCTCAGCGCGGGGATCTCATGCTGTGAGTTCTTGCCGCCCACTTGTTAT 1709
QY 1041 ----- 1040
DB 1710 TGACCTTATATGTTGTTACAATAAAGCAATAGCATCAAAATTTCAAAATTAAGCAT 1769
QY 1041 ----- 1040
DB 1770 TTTTCACTGCATTCATGTTGTTGTGTTGTTCCAACTCATCATGATGATCATGCTG 1829
QY 1041 ----- 1040
DB 1830 TATACCGTCGACCTCTAGCTAGAGCTTGCGTAAATGATGATGATGTTCTGTGTG 1889
QY 1041 ----- 1040
DB 1890 AAATGTTATCCGCTCACAAATTCACACACATACGACCGGAACATTAAGTGAAGC 1949
QY 1041 -----tgcggtgggtcgtactgcttct-- 1062
DB 1950 CTGGGTCCTAATGATGAGCTAATCACTCAATTAATGCTGTTGCGCTACGTGCCGCTTT 2009
QY 1063 -----gaagcggaagaacacagctgcatatgaatcgcccaagcgcggggagag 1114
DB 2010 CCAGTCGGGAAACCTGTGCTGCACCTCTCAATTAAGATTCGCCCAACGCGGGGAGAG 2069
QY 1115 cggttggtatctgggcgtcttcgcgtcctcgtgtaactgactcgtgcgtgcgtgcgt 1174
DB 2070 CGGTTGCGTATGCGCGCTCTTCGCTTCTCGCTCACTGACTGCTGCGCTGCGTGTGT 2129
QY 1175 tccgctcgcgcgcagcgtatcagctcactcaaaagcggtlaaaggttataccacagaatc 1234
DB 2130 TCGGCTGGGGAGGCGGATACGCTCACTCAAAAGCGGTTAATAGCTTATCCACGAATTC 2189
QY 1235 aggggatacgcaggaagaacatgtagcaaaaagccagcaaaaagccaggaacggttaa 1294

|||||
Db 2190 AGGGGATTAACGGCAAGAAACATGTAGCAAAAAGCCACAAAAGCCAGAAACCTTAA 2249
QY 1295 aaagccgctgtcgtgctgttccataagctccgccccctgacgagcatcaaaaa 1354
Db 2250 AAAGGCGCGTCTGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAA 2309
QY 1355 tcgagctcaagtcagaggtggtggaaccccgacagagctataaagatacaggcttcc 1414
Db 2310 TCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGACTATAAGATACAGGCGTTTCC 2369
QY 1415 ccttgaagctccctgtgctgtccctgttccgaacctccgcttaacgatacctgtc 1474
Db 2370 CCTGGAAGCTCCCTCGTGGCGCTCTCTGTCCGACCTCCGCTTACCGGATACCTGTC 2429
QY 1475 cgccttccctcctcggaaagcgtgagcgttcccaatgtcaccgctgtagatcag 1534
Db 2430 CGCCTTCTCCCTTCGGGAGCGTGGCGCTTTCATAGCTCAAGCTGTAGTATCTGAG 2489
QY 1535 ttgagtgtagtgcgttcgcctcaagctggtgtgtgtcagaacccccgttaagccga 1594
Db 2490 TTGCGGTGATAGTGTGCTGCTCAAGCTGGGCTGTGTGACAGAACCCCGTTCAGCCGA 2549
QY 1595 ccgctgaccttaccggttaactatcgtcttgaagtcacacccggttaagacagactatc 1654
Db 2550 CCGCTGGCGCTTATCCGTTAACTATCTCTTGAAGTCCAAACCGGTAAAGACAGACTTATC 2609
QY 1655 gccactgacagacacacactgttaacagagatlaagacagagatlatgtagcggtgtac 1714
Db 2610 GCCACTGGCAGACACCTGTAAACAGATTAGACAGACAGGATATAGGCGGTGTAC 2669
QY 1715 agagcttcaagtgtgtgaccttaactacggtctacagaagaataattgttactgt 1774
Db 2670 AGAGTTCTTAAAGTGAGGCTTAACAGGCTACACTAGAGAGGACAGTATTGTATCTG 2729
QY 1775 cgcctgctgaagcacaacttacccttcggaagaagttgtgtacgtcttgaaccgcaaca 1834
Db 2730 CGCTGTCTCAACGCACTTACCTTCGAAAAAGATTGTGATGCTTGTGATCCGCAACA 2789
QY 1835 aaccacgcgtgtgtagcgtgtgttcttctgttgcgaagcagagatlaagcgcagaaaa 1894
Db 2790 AACCAACCGCTGTAGCGGT 2849
QY 1895 agagatcaagaagatccttgcatttcttaaggggtcgcgaagctagtgtgaagaaaa 1954
Db 2850 AGGATCTCAAGAAGATCTTGTGATCTTGTACGGGCTGTGACGCTGAGGAAAGAAAA 2909
QY 1955 ctacagttagagattgtgcatgagcagatacatattgaatgattaga 2007
Db 2910 CTCACGTTAAGGATTTTGTGATGATGATTAACAAAAAGATCTTACCTAGA 2962

RESULT 3
US-08-760-615-7
Sequence 7, Application US/08760615
Patent No. 6200959
GENERAL INFORMATION:
APPLICANT: Haynes, Joel R
APPLICANT: Schmaljohn, Connie S
APPLICANT: Fuller, Deborah L
APPLICANT: Schmaljohn, Alan
APPLICANT: Jahrling, Peter B
TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229, 91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 4326 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Expression vector"
IMMEDIATE SOURCE:
CLONE: pMRG7077
FEATURE:
NAME/KEY: promoter
LOCATION: 1250..2062
FEATURE:
NAME/KEY: Intron
LOCATION: 2063..2887
OTHER INFORMATION: /function= "Human Cytomegalovirus"
FEATURE:
NAME/KEY: polyA-site
LOCATION: 2912..3314
FEATURE:
NAME/KEY: CDS
LOCATION: complement (299..1114)
US-08-760-615-7

Query Match 13.0%; Score 1053.6; DB 4; Length 4326;
Best Local Similarity 94.1%; Pred. No. 7, 6e-227;
Matches 118; Conservative 0; Mismatches 39; Indels 31; Gaps 1;

QY 820 tgatcaagcctcgaactgtgcttctgaattgccaagccatctgttgcctccctcccggtg 879
Db 3030 TGATCAGCCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTGTTTCCCTCCCGCGTG 3089
QY 880 ccttcccttgacctggaaggtgccaactccactgttcccttccataaaaaatgagaaat 939
Db 3090 CCTTCCCTTGAACCTGGAAGGTGCCACTGCCACTGCTCTTCCATTAATAATGAGAAAT 3149
QY 940 gcatcgcatgtctgaatgagtgatcattctatctctggtggtgtggtgtggtgagaaagc 999
Db 3150 GCATCGCATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3209
QY 1000 aaggggaagatttggaagaacaatagcagagcatatcgtggagatggtgtggtgtatggt 1059
Db 3210 AAGGGGAGATTGGAAGCAATAGCAGGCACTGCGGATGCGGTGCGCTCATAGCT 3269
QY 1060 tctgaagcggaaagaaacacagctgcatatgaatcggccaacggtcggtgagagcggtt 1119
Db 3270 TCTGAGCGGAAGAAACACACT-----GGGCTC 3298
QY 1120 tgcgtattggtgccttccgcttccctgcgtcaactgactgcgtggtgtgtgtgtgtgtgc 1179
Db 3299 GACAGCTCGACTCTGAATGCTTCCGCTCAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 3358
QY 1180 tgcggagacgtatagctcaactcaagcgggttaacggttaacagaaatcaaggg 1239

Db 3359 TCGGCGAGCGGTATACGCTACCTCAAGCGGTATACGATTACACAGATCAGGGG 3418
Qy 1240 ataagcaggaagacatgttagcaaaagccagcaaaagccaggaacggtlaaag 1299-
Db 3419 ATAAAGCAGAAAGACATGTAGCAAAAAGCCAGCAAAAAGCCAGAACCTTAATAAG 3478
Qy 1300 ccgctgtgctgttcttccatagctccgccccctgacagatcacaanaatcgac 1359
Db 3479 CCGCGTGTGCTGCTTTTCCATAGGCTCCGCCCTGACGAGATCCAAATAATCGAC 3538
Qy 1360 gctcaagtcagagtgaggaaacccagcagacataaagatacagagtggttccctcg 1419
Db 3559 GCTCAAGTCAGAGTGCGGAAACCCGACAGGACTATAAAGATACAGGGGTTTCCCTCG 3598
Qy 1420 gaagctccctcgtcgtctcctgttccagacctccgcttaccagatccgtccgact 1479
Db 3599 GAAGTCCCTCCTGCTGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCT 3658
Qy 1480 ttctccctcgggaagcgttgagcttctcctcaatgctcagcgttagtctcagttcgg 1539
Db 3659 TTCTCCCTTCGGGAAGCGTGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTGCG 3718
Qy 1540 tgaagtcgttgcctccaaagcgtggtgtgtgacgaaccccccgcttaagcccgact 1599
Db 3719 TGTAGTGTGCTGCTCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCGCGCT 3778
Qy 1600 gcgccttaccgtaactacgtcttgaagtcacacccggtlaagacacgactatccgac 1659
Db 3779 GCGCTTATCCGTAACATATGCTCTTGAAGTCCACCCGTTAAGACACGACTTATGCCAC 3838
Qy 1660 tggcagcagcactgtgtaacaggaattagcagagcagatgtagcagtgctacagagt 1719
Db 3839 TGGCAGCGACCACTGTAAACAGATAGCAGAGCAGAGATATAGCGGCTGTACAGAGT 3898
Qy 1720 tcttgaagtggtgagctacactacagctacaggaagaacatatttgatctcgcgc 1779
Db 3899 TCTTAAAGGGGGCTTACCTAGCTACACTAGAAAGAACATATTTGGTATCTCGGCTC 3958
Qy 1780 tgcctaaagcacttactcgcgaagaaggtgtgtagctctgtacccgcaacaacaa 1839
Db 3959 TGCTAAAGCAATTAACCTTGGGAAAAAGTGTAGCTTGTATCCGCAACAAACCA 4018
Qy 1840 ccgctgtgtagcgggtgttttttltgttgcgaagcagatcagcgcagaaaaaagat 1899
Db 4019 CCGCGGTAGCGGCTGTTTGTGTCAGACAGCAGATACGGCGCAAAAAAGGAT 4078
Qy 1900 ctcaagaagatccttgaatcttcttcaaggggtcgaagcgaagtggaagcaaacctac 1959
Db 4079 CTCAGAAAGATCTTGTGATCTTCTACGGGGTGTGAGCTGTAGTGAACGAAAACTC 4138
Qy 1960 gtaaggaattgtgcatagcagatacatatttgaatgtattaga 2007
Db 4139 GTTAAGGATTTTGTGTCATGATATATCAAAAAAGATCTTCACCTAGA 4186

RESULT 4
US-08-893-327-15
Sequence 15, Application US/08893327
Patent No. 6020192
GENERAL INFORMATION:
APPLICANT: Zolotukhin, Sergei
APPLICANT: Hauswirth, William W.
APPLICANT: Muzyczka, Nicholas
TITLE OF INVENTION: Humanized Green Fluorescent Protein
TITLE OF INVENTION: Genes and Methods
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,327
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UFLA.062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 6253 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 988..1701
US-08-893-327-15

Query Match 12.1%; Score 981.8; DB 3; Length 6253;
Best Local Similarity 86.9%; Pred. No. 1.1e-210;
Matches 1162; Conservative 0; Mismatches 37; Indels 138; Gaps 2;

Qy 809 ctgaagctcgtctgatacagctcgtcgtctgcttctagtgtccagccatctgttctgc 868
Db 3017 CTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCGAGCAATCTGTTTGGC 3076
Qy 869 cctcccccgtgctctcttgaacctggaaagtgccaccccacttcttccctaataa 928
Db 3077 CTTCCCGGTGCTTCTTGAACCTGGAAGTGCCACTCCACTCTCTTCTTATATAA 3136
Qy 929 atgaagaaatgcatcagatgtctgaagtaggtgtgcatctatcttctggggggtgg 988
Db 3137 ATGAGGAATGTGATCGATGTGTGTGAGTAGTGTGATTTCTTGGGGGGTGGGG 3196
Qy 988 ggcaggaacgcaagggggaagatttggaaagacaataagcagcagtctggga 1040
Db 3197 GGCAGGACGCAAGGGGGGAGATTGGAAAGACATAGCAGCATCTGGGAGAGATCTA 3256
Qy 1041 ----- 1040
Db 3257 GGAACCCCTAGATGATGAGTTGGCCACTCCCTCTGCGCGCTCGCTGATCAGAGC 3316
Qy 1041 -----tgcggtgggtctctatgtcttctgaagcgggaagaaacag-- 1079
Db 3317 GCGCGGGGCAAGCCCGGGGCGCGGGCGACCTTGTGTGCGCGGCTCAGTGAGGAGCG 3376
Qy 1080 -----ctgataatg 1090
Db 3377 AGCGGCGAGAGAGAGTGCGCAACCCCGCCCGCCCGCTGAGCGCTGCAATTAATG 3436
Qy 1091 aatggcaaacgscgggggaaggggttggtaatttgggcgcttccgttccgtctgct 1150
Db 3437 AATCGGCAACGCGGGGAGAGCGGTTGTGATTTGGGCGCTTCCGCTTCTCTCCT 3496
Qy 1151 cactgaactcgtcgtcgtctgtctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1210
Db 3497 CACTGACTGCGTGGCTGT 3556
Qy 1211 gtaatacgttatacacaagaatcaggagataacgcaagaaacatgtgagcaaaag 1270

Db 3584 GGTAAATACGGTTATCCACAGATACGGGATACGACAGAAAGACATGTGACGAAAAAG 3643
Qy 1271 ccagcaaaagccaggaacccgtaaaagccgctgtgctgctgttccatagctcgc 1330
Db 3644 CCACGAAAAAGGCCAGAACCTTAATAAGGCCGCTGTGCTGGCGTTTTCATAGCTCCG 3703
Qy 1331 ccccccctgaagacatcaaaaatcgacgctcaagtcagagtgagcgaacccagcag 1390
Db 3704 CCCCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGCGCAAAACCCGACAG 3763
Qy 1391 actataagataccaggggtttcccccctggagagctccctgtgctgtctctgttcgac 1450
Db 3764 ACTATAAAGATACCAAGGCGTTTCCCTGGAAGCTCCCTGTGCGCTCTCCGTTCGAC 3823
Qy 1451 cctgcgcctacccgataccctgtccgccttccctccctcggaaacgctgagccttcca 1510
Db 3824 CCTGCGGTTACCGGATACCTGCTCGCTTCTCCTTCCGGGAACGCTGGCGCTTTCAC 3883
Qy 1511 atgtcagcgtgtatgtatctcaagttcggtgtaggtcgctgcctcgaagctggcgtgt 1570
Db 3884 ATGCTACGCTGTAGGTATCTCACTTCGCTAGGTGCTGCTGCTCAAGCTGGGCTGTGT 3943
Qy 1571 gcaagaaaccccgcttcaagcccgccgctgctgcttaccggttaactatcgtcttgatc 1630
Db 3944 GCACGAACCCCGTTCAGCCCGACCGCTGCGCTTATCCGTAACATVCCCTTTGAGTCC 4003
Qy 1631 caaccgagtaagacagactatcgcacactgacagcaacacgtaacagattagcag 1690
Db 4004 CAACCCGGTAAAGACACACTTATCGCCACTGACAGCAACCACTGGTAACAGATTAGCAG 4063
Qy 1691 agcagagtatgtagcggtgtctacagagttcttgaagtgtgtgccttaactacgcttacac 1750
Db 4064 AGCAGAGTATGAGCGCGTGTCTACAGAGTTCTGAAGTGCGCTTAACCTACGCGCTACAC 4123
Qy 1751 tagaagagcagatttggtatctgctgctgctgctgaagcaattaccttcggaaaaaggt 1810
Db 4124 TAGAAGGACATATTTGGTATCTCGCTGTCTGTGAAGCACTTACCTTCGGAAGAAAGAGT 4183
Qy 1811 tggtagctctgtatcgcgcaacaacaacacgcgtgtagcggtgtttttgtttgtaa 1870
Db 4184 TGGTAGCTCTTGATCCGGCAAAACACACGCGTGTAGCGGTGTTTGTGTTTGCAA 4243
Qy 1871 gcagcagattacgcgcagaaaaaagaatcacaagaagatccttgcattcttctacgag 1930
Db 4244 GCACGACATTCACGCGCAGAAAAAGATCTCAAGAAATCTTTGATCTTTCTACGCG 4303
Qy 1931 gtcgaagcctcagtggaagcaaaactcacgttaaggaattgtgtcatagacgatacat 1990
Db 4304 GTCTGACGCTAGTGGAACGAAACTCACGTTAAGGGAATTTGGTCATAGGATTAATCAAA 4363
Qy 1991 attgaatgattaga 2007
Db 4364 AAGGATCTTCACCTAGA 4380

RESULT 6
US-08-893-327-19
; Sequence 19, Application US/08893327
; Patent No. 6020192

GENERAL INFORMATION:
; APPLICANT: Zoelotukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,327
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/588,201
; FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: DPLA:062\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 988..1728
; US-08-893-327-19

Query Match 12.1%; Score 981.8; DB 3; Length 6280;
Best Local Similarity 86.9%; Pred. No. 1.1e-210;
Matches 1162; Conservative 0; Mismatches 37; Indels 138; Gaps 2;

Qy 809 ctagagctcgtatgatacagctcgtatgcttctatgttgcacgacatctgtttgctc 868
Db 3044 CTAGAGCTCGCGTATACACCTCGACTGCTTCTAGATTGACAGCACTGTGTTGCC 3103
Qy 869 cctcccgcgttcttcccttgcacccctggaaagtgccacccacactgcttccctaataa 928
Db 3104 CTTCCCGGTGCTTCTTGAACCTGGAAGGTGCCACTGCCAGTCTCTTCTCTAATAA 3163
Qy 929 atgagaaatgtacatcagatgtctgtagagtgatcattctatctcgtgggggtgggg 988
Db 3164 ATGAGGAATTCATACGATGTTGTGATAGGTGATTCATTCTGGGGGTGGGGTGG 3223
Qy 989 ggcagagacagcaaggggagagatgtgggaagaacaatagcagcatgtctggga----- 1040
Db 3224 GGCAGGACAGCAAGGGGAGGATGGGAAGCAATAGCAGGCAATCTGTGGGAGAGATCTA 3283
Qy 1041 ----- 1040
Db 3284 GGAACTCTAGTATGATGATTGGCCACTCCCTCTCTGCGCGCTGCTGCTCATAGTACG 3343
Qy 1041 -----tgcygtggctctatgtcttctgagcggaagaaacag-- 1079
Db 3344 CGCCCGGCAAAAGCCGGGGGCTCGGGGACCTTTGTGTGCGCCGGCTCAAGTGAAGGACG 3403
Qy 1080 -----ctgacattaag 1090
Db 3404 AGCGCGCAGAGAGGAGTGGCCAAACCCCGCCCGCCCGCCCGCTGACGCTGCAATTAATG 3463
Qy 1091 aatgggcaaaagcgcggggaaagcggtttgctgatatggggtcttccggttctcgtc 1150
Db 3464 AATCGCCAAACGCGCGGGGAAGGCGTTTGCATTATGGGCGCTTCTCCCTCGCT 3523
Qy 1151 cactgaactcgtcgcgtggtcgttctgagctgagcgagcgtlatcagctcaaaagc 1210
Db 3524 CACTGACTCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3583
Qy 1211 gtaatacgttatccacagaatcaggggataacgcaaggaagacatgtgagcaaaag 1270
|||||

Db 3584 GGTAAATACGGTATTCACAGATCAGGGATTAACGAGAAAGACATGTGACAAAGG 3643
QY 1271 ccagcaaaagccaggaacccgttaaaagccggttgctgctgttccataggtccg 1330
Db 3644 CCAGCAAAAGCCAGGAAACCGTAAAAAGCCGGTGTGCTGCTTTTTCATAGGCTCCG 3703
QY 1331 ccccccctgaacgacatcacaaaatcgaagctcaagtcagaggttgaggaaccgagag 1390
Db 3704 CCCCCCTGAGAGACATCACAAAATGACGCTTAAGTCAGAGGTGGCGAACCAGACAGG 3763
QY 1391 actataaagatacagagcggttcccccctggaagctccctcgtgcgtctcgttccgac 1450
Db 3764 ACTATAAAGATACAGAGCGGTTTCCCTGGAGCTCCCTGCGGCTCTGTTCCGAC 3823
QY 1451 cctgcgcgttacggaatacctgtcgcgcttctcccttcgggaagcgctggcgttctca 1510
Db 3824 CCGCGCGCTTACCGGATACCTGTCCGCTTTCCTTCGGAAGCGTGCGCTTCTCA 3883
QY 1511 atgtcaagctgtgaagtaatcagttcgtgtgagtcgttcggtcgaagctggcggtgt 1570
Db 3884 ATGCTACACGCTGTAGATCTCAAGTTCGGTGTAGTGTGCTTCCAGGCTGGGCTGTGT 3943
QY 1571 gcaagcaaccccccgttcagcccgacgctgagccttaccgtaactatcgtctgagtc 1630
Db 3944 GCACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGTAAGTATGCTTTGAGTC 4003
QY 1631 caaccggttaagacagcaactatcgcgaactgagcagcaactggttaacagattagcag 1690
Db 4004 CAACCCGGTAAGACACAGACTTATCGCCACTGCGACAGCACCTGTAACAGGATTAGCAG 4063
QY 1691 agcgaagctgtgtagcggtgctcagagttcttgaagttgtagcctaactagcgctaac 1750
Db 4064 AGCGAGGTTATGTAGCGGGTGTCTACAGATTTCTTGAAGTGGTGGCTTACCTACGCTAC 4123
QY 1751 tagaaggaacagtaattgtatctcgtcgtcgtcgaagcgaactaccttcggaaaagag 1810
Db 4124 TAGAAGGACAGTATTGTGATCTGCGCTGTGTAAGCCAGTTACCTCGGAAGAAAGAGT 4183
QY 1811 tggtagctttagtcggaacaaacacacgctgtagcggtgtgttttttttggaa 1870
Db 4184 TGGTACTCTTGTATCCGCGAAMACACACCGCTGAGCGGTGTTTGTGTTGCAA 4243
QY 1871 ggaagcagattacgagagaaaaaagatctcagaagatcccttatcttcttcgag 1930
Db 4244 GCAAGCATTTACCGCGAAGAAAGAGATCTCAGAAAGATCTTGAATCTTCTACGGG 4303
QY 1931 gtcagcgtcagtggaacgaacacacgctlaaggaatttgatcatgagcgatacat 1990
Db 4304 GTCGTAGCGCTCAAGTGAACGAACATCAGCTTAAGGATTTGTGATGAGATTATCAAA 4363
QY 1991 attgaatgatttaga 2007
Db 4364 AAGGATCTTCACCTAGA 4380

RESULT 7

US-08-343-401A-3
; Sequence 3, Application US/08343401A
; Patent No. 5661132
; GENERAL INFORMATION:
; APPLICANT: Swain, William F
; APPLICANT: Macklin, Michael D
; APPLICANT: Eriksson, Eilof
; APPLICANT: Andree, Christophe
; TITLE OF INVENTION: Improved Wound Healing
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: PO Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,401A
FILING DATE: 22-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 11-229-9103-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: PMR01630
FEATURE:
NAME/KEY: exon
LOCATION: 713..721
FEATURE:
NAME/KEY: exon
LOCATION: 981..1253
FEATURE:
NAME/KEY: CDS
LOCATION: join(713..721, 981..1253)
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 713..1049
US-08-343-401A-3
Query Match 12.0%; Score 973; DB 1; Length 4283;
Best Local Similarity 83.2%; Pred. No. 8.6e-209;
Matches 1216; Conservative 0; Mismatches 30; Indels 215; Gaps 1;
QY 762 ctgagacatcactatgagggccctattctatagtgtaacctaaatgaggaattgc 821
Db 1277 CTCGAGCATCATCTAGAGGGCCCTATTCTATAGTGTACCTTAATGCTAGAGCTGCTG 1336
QY 822 atgaagctgactgtgctctcttagttgcaagcagcatcgtgtgttgcctcccccgtgc 881
Db 1337 ATCAGCCTGACATGTCCTTCTAGTTCAGCCATCTGTTGTTGCCCTCCCGTGC 1396
QY 882 ttccctgaacctggaaggtgacacccactgctcccttccctaaatgaggaattgc 941
Db 1397 TTCTTGACCTGTGAAGGTGCTCCACTGCTCTTCTTAATTAATAGGAATTC 1456
QY 942 atcgcatgtctgtagtagtgatcattctatctcgtggtggtggtggtggtggtggtggt 1001
Db 1457 ATCGCATTTGCTGAGTAGGATGATCTTATTTGGGGGGTGGGGGGTGGGGGACGACAA 1516
QY 1002 gggggaagattggaagacaatagcaggaacgtctggggatggtgtggtctatgcttc 1061
Db 1517 GGGGGAAGATTGGGAAGACAATAGCAGGATGCTGGGGATGCGTGGGCTCTATGGAAC 1576
QY 1062 tgagcggaaga----- 1074
Db 1577 AGCTGGGGCTCGAGATGCAAGCTTGAATTTCTAATAGTGTACCTTAATAGCTTGCGCT 1636
QY 1075 ----- 1074
Db 1637 AATCATGTCATAGCTGTTTCTGTGGAATTTGTTATCCGCTCACAATTCCACACACA 1696
QY 1075 ----- 1074

[illegible]

```

US-08-445-265A-1
; Sequence 1, Application US/08445265A
; Patent No. 5697901
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Elof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,265A
; FILING DATE:
; CLASSIFICATION: 604
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229.91080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-9166
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
US-08-445-265A-1

```

Query Match	12.0%	Score 973:	DB 1:	Length 4283:
Best Local Similarity	83.2%	Pred. No. 8.6e-209:		
Matches 1216:	Conservative	0:	Mismatches 309:	Indels 215: Gaps 1:
QY	762	ctcgcgcattgcatactagaaggccctattctatagcgttaaccctaattgctagagtcgcgtg	821	
Db	1277	CTCGAGCATGCTATGAGAGGCGCCATTCTAATAGTCACTAAATGCTAGAGCTCGCGTG	1336	
QY	822	atcaaaccttcgaatctgycctcttaagtcccaagcatctgtttgcccctcccgcgtgc	881	
Db	1337	ATCAGCCTCGAATGtgccttCTTAGTTGCCAGGCATCTGTTGTTGCCCTCCCGCGTGC	1396	
QY	882	ttcccttgaccctcggaagcgtgcaccctcccatctgccttccctaataaalagagaaattgc	941	
Db	1397	TTTCCTTGACCCCTGGGAAGGCGCACCTCCACGTCTCTTTCCTAATAATGAGAAATTGC	1456	
QY	942	atcgcattctcgtatagtgatcatctatctcttggggggttggggttcgagcaagcaaa	1001	
Db	1457	ATCGCATTTCTCGAATAGTGTCACTTCAATTCTGGGGGTTGGGGGAGGAGCAAGCA	1516	
QY	1002	gggggagagatctggaagaacaatagcaggcatctgtgggagtgcgttgggctctatgcttc	1061	
Db	1517	GGGGGAGGATTTGGGAAGAACATAGCAAGCATCTGGGAGTGGGCTGTATGGAAAC	1576	
QY	1062	tgaagcgygaaga	1074	
Db	1577	ACCTGGGGCTGCAGATGCCAAGCTTGAGATTCTATAGTGCACCTAATAATAGCTTGGCGT	1636	
QY	1075		1074	

D	b	1637	AATCATGGTCTACTACGTCTTTCCGTCGTGTAAGAAATTGTTATCCCGCTCAACAATTCCACACANCA	1636
O	y	1075	-----	1074
D	b	1697	TACGAGCGGAAGCATAAAGTGTAAGAACCAGGGGTCCTAATGAGTAGACTAACTCACAT	17566
O	y	1075	-----accagctgcatt	1086
D	b	1757	TAAATTGCATTGCGCTCACGTCCCGCTTTCAGTCCGGAAAACCTGTCGTGCCACAGCTGCATTT	1816
O	y	1087	aatgatcgcgcaacgcgcggaggagaagcggttttgtaatttggcgcccttcocgcttcct	1146
D	b	1817	AATGATATGGCCAAACGGCGGGGAGAGCGGTTGGGTAFTTGGCGCCTTCCGCTTCT	1876
O	y	1147	cgcatacgtactcgctcgctcgctgcgttgcgcgtcgcgagcgytatcaagtcaactaa	1206
D	b	1877	CGCTCACTAGACTCCGCTCGCTCGGTGGTTCGGCTCGCGGACGGGTATCAGCTCACTCAA	1936
O	y	1207	aggcgttaatacgtttatccaagaattccagggtatacgcgaagaaagaaactgtgagcaa	1286
D	b	1937	AGGGCGTAATAACGGTTATCCACAGAAATCAGGGGATATACCCAGAAAGAACATGTACGAA	1996
O	y	1267	aaggcgcagcaaaaagcgcaaggaaacgcgtlaaaaaagcgcgcttgtctgcgqtlttlccaatgc	1326
D	b	1997	AAGGCAACACAAAGCCAGGAGACCGTTAAAAAGGCGCGCTTGCTGGCGTTTTTCGATTAAGC	2036
O	y	1327	tccgcgcccccctgacgagcatcacaaaaatlcgacyctcaagtacaaggtgtgcgaaccga	1386
D	b	2057	TCCGCCCCCTCGACGAGCATTCACAAAAATCGACGCTCAAGTACAGAGTGCGGCAAAACCGGA	2116
O	y	1387	caggagctaataagaataccaagcgcttccccctbgaaagtcacctcgtagcctccctgtc	1446
D	b	2117	CAGGACTTATAAGATACCAAGGCGTTTCCCCTGGAAGCTCCCTCGGGCCTCTCCTGTTCC	2176
O	y	1447	cgaccctgcgcgttacaccggaatacctgtlcgcgccttltcccttcgggaaagcgtgtgcgctt	1506
D	b	2177	CGACCCCTGCGCTTACCGSGATACCTGHCGCCCTTTCCTTCCTTCGGGAAGCGTGGCGCTTT	2236
O	y	1507	ctcaatgtctcaagcgtctagtatatactcaagttcgggtgtaggtgtgtgcgtccaaagctgggct	1566
D	b	2237	CTCAATAGCTCTACGCTGTATAGGTATCTCAGTTCGGGTGTATGGTCTGCTCCAAAGCTGGGCT	2296
O	y	1567	gtgtgcagcaaaccccccgcttcacgcgcgaaccgcgctgtgccttaccggtiaactatcgtcttg	1626
D	b	2297	GTTGTCAAGAACCCCCGTTTACGCCAGCCGCTGCCCTTATCCGTTACTATACGTCCTTGG	2356
O	y	1627	agtcacaacccggtlaagacaacagacttalcgcacattgycacgacgacacactgtglacagatat	1686
D	b	2357	AGTCCAACCCGCTTAAGACACAGACTTATCCGCATATGGCAGACAGCACTGGTAAACAGATTA	2416
O	y	1687	gcagaagcaggttatgtatagcggtgtactacagagttcctttaagtgtgtgcctaaactaagcgt	1746
D	b	2417	GCAGGACAGAGTATGTAGCGGTGCTACAAAGATTCTTGAAGAGTGTCGCCCTTAACCTACGCT	2476
O	y	1747	aacacagaagaagcagtaatttggfatactgcgcctctgtctgaagccadttacccttcggaaaaa	1806
D	b	2477	ACACTAGAAAGACAGTATTGTTGGTATCTGCCCTGTGCTGAAGCCAGTTACCTTCCGAAAAA	2536
O	y	1807	gaattgtagttactcttgatccgcgaacaacacccgcgtgtglatcgytggatttlltctt	1866
D	b	2537	GAGTTGGTAGCTTGTGATCCGGCAACAAACACCCGCTGTATGCGGTGATTTTTTTGTTT	2596
O	y	1867	gcaagcagcagatatacgcgcgaagaaaaaagagatctcaagaagaatccctttgatcctttcta	1926
D	b	2597	GCAAACACACATATTACGCGCGAAAAAAGATCTCAAAACATCCCTTGTATCTTTTCTTA	2656
O	y	1927	cggagtcgcagcctcaatgtgaaagaaaaccaactcaagtaaagatttggatcagagcgat	1986
D	b	2657	CGGGGTCTGACGCTCACTGTGAACAGCAAAACCTCACGCTTAAAGGATTTTGGTCATGAGATTAAT	2716
O	y	1987	acataattgaatgtattataga 2007	

```

Db      2717  CAAAAGATCTTCACCTGA  2737

RESULT 9
US-08-990-442-1
: Sequence 1, Application US/08990442
: Patent No. 6090790
: GENERAL INFORMATION:
: APPLICANT: Eriksson, Elof
: TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Quarles & Brady
: STREET: 1 South Plinckney Street
: CITY: Madison
: STATE: WI
: COUNTRY: US
: ZIP: 53703
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/990,442
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Berson, Bennett J
: REGISTRATION NUMBER: 37094
: REFERENCE/DOCKET NUMBER: 310558.90028
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 608-251-5000
: TELEFAX: 608-251-9166
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4283 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "Plasmid DNA"
: FEATURE:
: NAME/KEY: CDS
: LOCATION: join(713..721, 981..1250)
US-08-990-442-1

Query Match      12.0%; Score 973; DB 3; Length 4283;
Best Local Similarity 83.2%; Pred. No. 8.6e-209;
Matches 1216; Conservative 0; Mismatches 30; Indels 215; Gaps 1.

QY      762  ctcgagcatgcatctcagagggccctattctataagtgtcaaccctaaatgctagagctgcgtg  821
      |||||||
Db      1277  CTCGAGCAATCATCAGAGGGCCCTATTCTATAGTCACTAAATGCTAGAGCTCGCTG  1336
      |||||||

QY      822  atcaacccctgaagatgtgaccttaagtgtgcaagcatctgtgtgtgccctcccccgtgac  881
      |||||||
Db      1337  ATCAAGCTTCGATGTGCTCTTCTAGTTGCCAGGCACTGTGTTGGCTCCCTCCCGCTGCC  1396
      |||||||

QY      882  ttccttgaacctggaagtggtccactccactgttccttctcctaataatgagaaatgtgc  941
      |||||||
Db      1397  TTCCTTGAACCTCGAAGGAGGTCACATCCCACTGCTCTTCTTAATAAATGAGGAATTGC  1456
      |||||||

QY      942  atcgcatgtctcgaataggtgtcattctatcttcctgggggtgggtggggcaagacaagaa  1001
      |||||||
Db      1457  ATCGCATGTCTCGAATAGGTGTCAATTTGTGGGGGCTGGGGGAGAGACAGCA  1516
      |||||||

QY      1002  gggggagagatggggagagacaatagcagagcatgctggggaatgagtgaggtctctatgcttc  1061
      |||||||
Db      1517  GGGGAGAGATTGGGAAGACAATAGCAGGCACTGTGGGAGTGGGCTCTATGGAAAC  1576
      |||||||

QY      1062  tgaagcggaagaa-----  1074

```



```
Db 1577 AGCTGGGCTCGAGCAGTCAAGCTTGAGTATTCTATAGTGTACCTTAATAGCTTGCGGT 1636
Oy 1075 -----
Db 1637 AATCATGTGCTATAGCTGTTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCACACACACA 1696
Oy 1075 -----
Db 1697 TAGCAGCCGGAAGCATAAAGTAAAGCTGGGGTGCCTAATGAGTAGATACATACAT 1756
Oy 1075 -----
Db 1757 TAAATTGCGTTCGCTCACTGCCCCGCTTTCAGTCGGGAACCTGTGTCGACAGCTGCATT 1816
Oy 1087 AATGAATCGCAACGCGCGGGAGAGCGGTTTGCGTATTGGGGCGCTCTTCCGCTTCT 1876
Db 1817 AATGAATCGCAACGCGCGGGAGAGCGGTTTGCGTATTGGGGCGCTCTTCCGCTTCT 1876
Oy 1147 CGCTCACTGACTCGTCTGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
Db 1877 CGCTCACTGACTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1936
Oy 1207 AAGCGTGAATACGTTATCCACAGATCAGCGGATTAACGCAAGAAACATGTCAGCAA 1266
Db 1937 AAGCGGTAAATCGGTTATCCACAGATCAGCGGATTAACGCAAGAAACATGTCAGCAA 1996
Oy 1267 AAGCGCAAGAAAGCGCAAGAAAGCGGATTAAGCGGATTAAGCGGATTAAGCGGATTAAGCG 1326
Db 1997 AAGCGCAAGAAAGCGCAAGAAAGCGGATTAAGCGGATTAAGCGGATTAAGCGGATTAAGCG 2056
Oy 1327 TCCGCCCCCTGACGAGATACACAAGAAATCGAGCTCAAGTCAAGAGTGGGAAACCCGA 1386
Db 2057 TCCGCCCCCTGACGAGAGATACACAAGAAATCGAGCTCAAGTCAAGAGTGGGAAACCCGA 2116
Oy 1387 CAGGACTAATAAGATACAGAGCGGTTTCCCTGGAAGCTCCCTGCGCTCTCTCTGTTTC 2176
Db 2117 CAGGACTAATAAGATACAGAGCGGTTTCCCTGGAAGCTCCCTGCGCTCTCTCTGTTTC 2176
Oy 1447 CGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1506
Db 2177 CGACCCCTGCGCTTACCGGATACCTGTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2236
Oy 1507 CTCAATGCTCAAGCTGTAAGTATCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
Db 2237 CTCAATGCTCAAGCTGTAAGTATCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2296
Oy 1567 GTGTGCAAGAACCCCTTACGCCGACCCCTGCGCTTATCCGCTTATCCGCTTATCCGCTTAT 1626
Db 2297 GTGTGCAAGAACCCCTTACGCCGACCCCTGCGCTTATCCGCTTATCCGCTTATCCGCTTAT 2356
Oy 1627 AGTCAACCCGCTTACGACACGATTAATGCACTGACGACGACGACGACGACGACGACGACGAC 1686
Db 2357 AGTCAACCCGCTTACGACACGATTAATGCACTGACGACGACGACGACGACGACGACGACGAC 2416
Oy 1687 GCAAGAGCGAGTATGAGCGGTGTGTACAGAGTCTTGAAGTGTGCGCTAATCGGCT 2476
Db 2417 GCAAGAGCGAGTATGAGCGGTGTGTACAGAGTCTTGAAGTGTGCGCTAATCGGCT 2476
Oy 1747 ACACTAGAGGACAGTATTTGTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806
Db 2477 ACACTAGAGGACAGTATTTGTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2536
Oy 1807 GAGTTGTAGTCTCTTGTATCCGCAACCAACCGCTGTAAGCGGTGTTTGTGTTT 1866
Db 2537 GAGTTGTAGTCTCTTGTATCCGCAACCAACCGCTGTAAGCGGTGTTTGTGTTT 2596
Oy 1867 GCAAGCAAGATTAAGCGCGCAAGAAAGATCTCAAGAGATCTTGTGATCTTTTCTA 1926
Db 2597 GCAAGCAAGATTAAGCGCGCAAGAAAGATCTCAAGAGATCTTGTGATCTTTTCTA 2656
Oy 1927 CAGGAGTCTGACGCTGAGTGAAGCAAGAACTCAAGTGAAGGATTTGTGATGAGCGAT 1986
Db 1927 CAGGAGTCTGACGCTGAGTGAAGCAAGAACTCAAGTGAAGGATTTGTGATGAGCGAT 1986
```

```
Db 2657 CGGGCTCTGACGCTCAGTGAAGCAAGAAACCTCACGTTAAGGATTTTGGTCAATGAGATTAT 2716
Oy 1987 ACAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2716
Db 2717 CAAAAGATCTTTCACCTTGA 2737

RESULT 10
US-08-467-420A-49
: Sequence 49, Application US/08467420A
: Patent No. 5683892
: GENERAL INFORMATION:
: APPLICANT: Ames, Robert S.
: APPLICANT: Appelbaum, Edward R.
: APPLICANT: Chaiken, Irwin M.
: APPLICANT: Cook, Richard M.
: APPLICANT: Gross, Mitchell S.
: APPLICANT: Holmes, Stephen D.
: APPLICANT: McMillan, Lynette J.
: TITLE OF INVENTION: Recombinant IL5 Antagonists Useful In
: TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
: NUMBER OF SEQUENCES: 74
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: SmltnKline Beecham Corp./Corporate
: ADDRESSEE: Intellectual Property
: STREET: P. O. Box 1539-UW2220
: CITY: King of Prussia
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,420A
: FILING DATE:
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/363131
: FILING DATE: 23-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sulton, Jeffrey A.
: REGISTRATION NUMBER: 34,028
: REFERENCE/DOCKET NUMBER: P50282
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610 270-5024
: TELEFAX: 610 270-5090
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6285 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: circular
: MOLECULE TYPE: DNA (genomic)
: US-08-467-420A-49

Query Match 11.4%; Score 921.8; DB 1; Length 6285;
Best Local Similarity 82.9%; Pred. No. 2.9e-197;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;
```


|||||
Db 3829 GATGAGCCTGACGTGCTTCTAGTGTCCAGCCATCTGTGTGCCCCCTCCCGCTGC 3888
Qy 881 ctctcttacccttggaaggtgcacatccaccctgtctcttcccaataaatggaagtatg 940
Db 3889 CTTCCTTACCCCTGGAAAGGTGCACCTCCACCTGCTTCCATAAATAAGAGAAATTTG 3948
Qy 941 catgcattctctgtgtgagtgatcatctatctctgagggtggtggtgagcagca 1000
Db 3949 CATGCCATTGCTGTAGTAGTGTCTATCTATTCTGGGGGTGGGGTGGGGCAGACAGA 4008
Qy 1001 agggaggagattgggaagacaatagcagcagcagcagcagcagcagcagcagcagc 1060
Db 4009 AGGGGAGGATTGGGAGACATAGCAGGCATGTGGGATGCGGTGGGCTCTATGAAAC 4068
Qy 1061 ctgagcgggaaga----- 1074
Db 4069 CAGCTGGGGCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATG 4128
Qy 1075 ----- 1074
Db 4129 TCATGCTATAGCTGTTCTCTGTTGTAATTTGTATCCGCTCACATTCACACACATA 4188
Qy 1075 ----- 1074
Db 4189 CGAGCGGAGACATTAAGTGAAGCTGGGGGTGCTAATGATGATGATGATGATGATG 4248
Qy 1075 ----- 1074
Db 4249 ATTGGTTGCGCTCACTGCGCCCTTTCCAGTGGGAGAAACCTGTGCTGCTGATGATTA 4308
Qy 1089 tgaatcggccaaagcggcgggaggaggagcgttgcgtatgtggcgtctctctctctc 1148
Db 4309 TGAATCGGCGCAACGGCGGGAGAGCGGTGCGGATTTGGGCGGCTCTCGCTTCTCTG 4368
Qy 1149 ctctactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1208
Db 4369 CTCACGTGACTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 4428
Qy 1209 gggggaataaggttttcccaaatcaggggataagcgggaaagaaacatgtgagcaaa 1268
Db 4429 GGGGAATAAGGTTTATCCACGAATCAGGGGATTAACGGCAAGAAAGATGTGAGAAAA 4488
Qy 1269 gggcagcaaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1328
Db 4489 GGCCAGCAAAAAGCCAGACACCTGTAATAAAGCCGCTTCTGCGTTTCCATAGCTC 4548
Qy 1329 cggcccttgacgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1388
Db 4549 CGCCCTCGTACGAGCAGATCACAATAATCAGCCTCAAGTCAGAGGTGGGAAACCGACA 4608
Qy 1389 ggaactataagatlacagcggttctcccttggaagtcctcctcgtgcctcctcgtt 1448
Db 4609 GGACTATAAAGATACAGCGCTTCCCTCGGAAAGCTCCCTGCTGCTCTCTGTTCCG 4668
Qy 1449 accctgcgcgttaccggatcagcagcagcagcagcagcagcagcagcagcagcagc 1508
Db 4669 ACCCTGCCCTTACCGGATACCTGTCGCCCTTCTCTCTCCGGAAGCCTGGCGCTTCT 4728
Qy 1509 caatgctcagcgttgaagatctcagtcgttgcgttgcgttgcgttgcgttgcgttgc 1568
Db 4729 CAATGCTCAGCGTGAATCTCAAGTTGGGTGAGGTGTTGCGTCCAAAGCTGGCTGT 4788
Qy 1569 gtgcagaaaccccgcttaccgcgagcagcagcagcagcagcagcagcagcagcagc 1628
Db 4789 GTGCAGAAACCCCGCTTACGCCCGACCCGCTCGCTCTATCCGGTAACTATGCTTTGAG 4848
Qy 1629 tccaaaccggtgaagacagcagcagcagcagcagcagcagcagcagcagcagcagcag 1688
Db 4849 TCCAAACCGGTGAAGACAGACTTATGCGCAGTGGCGAGCAGCACTGTGTAAAGGATTAC 4908
Qy 1689 agagcaggtatgtagcgttgcacagagcttcttgaatgtgtggtcctaactagcgtac 1748
|||||

Db 4909 AGAGCAGATATGATGAGCGGTGCTACAGAGTTCCTGAAGTGTGCTTAACCTACGCTAC 4968
Qy 1749 actagaagacagatattggtatctgcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgt 1808
Db 4969 ACTAGAAGCAGATATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 5028
Qy 1809 gtgtgagctcttgaatccggcaacaacacacacacacacacacacacacacacacacac 1868
Db 5029 GTTGTAGCTTGTGATCGGCAACCAACACCCGCTGATGCGGTGTTGTTGTTGTTGTTG 5088
Qy 1869 aagcagcagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1928
Db 5089 AAGCAGCAGATTAACCGCAGAAAAAGAGATCTCAAGAAAGATCTTGAATCTTCTACG 5148
Qy 1929 ggtctgacgctcagtggaacgaacacacacacacacacacacacacacacacacacac 1988
Db 5149 GGTCTGACGCTCAGTGTGACCAAAACCTCAGCTTAAGGATTTGCTATGATGATTATCA 5208
Qy 1989 atattgaatgattaga 2007
Db 5209 AAAAGATCTTCACTAGA 5227
RESULT 12
US-08-667-769A-49
: Sequence 49, Application US/08667769A
: Patent No. 5783184
: GENERAL INFORMATION:
: APPLICANT: Ames, Robert S.
: APPLICANT: Appelbaum, Edward R.
: APPLICANT: Chaiken, Irwin M.
: APPLICANT: Cook, Richard M.
: APPLICANT: Gross, Mitchell S.
: APPLICANT: Holmes, Stephen D.
: APPLICANT: McMillan, Lynette J.
: APPLICANT: Theisen, Timothy W.
: TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
: NUMBER OF INVENTIONS: Treatment of IL5 Mediated Disorders
: NUMBER OF SEQUENCES: 76
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corp./Corporate
: STREET: P.O. Box 1539-DM2220
: CITY: King of Prussia
: STATE: Pennsylvania
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/667,769A
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/17082
: FILING DATE: 22-DEC-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/467420
: FILING DATE: 06-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/470110
: FILING DATE: 06-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/363131
: FILING DATE: 23-DEC-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sutton, Jeffrey A.
: REGISTRATION NUMBER: 34,028
: REFERENCE/DOCKET NUMBER: P50503
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5024

TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 6285 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-667-769A-49

Query Match 11.4%; Score 921.8; DB 1; Length 6285;
Best Local Similarity 82.9%; Pred. No. 2.9e-197;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;

QY 821 gacagcctgactgctgcttccttcagttgccaagcactgtgtgttgcctcccccgtgc 880
DB 3829 GATCAGCCTGAGCTGCTCTTCTAGTTGCGACCATCTGTTGTTGCCCTCCCTCCGTC 3888
QY 881 ctccctgacccctggaagtgccactcccactgtcccttccctaataatgaagaaattg 940
DB 3889 CTTCCTTGACCTTGGAAGGTGCCACTGCCACTGCTCTTCTTAATAAATGAGAAATTG 3948
QY 941 caccgcattgtctgaagtagtgatcattctatctctgagggttgagggtgagagacagca 1000
DB 3949 CATCGCATGCTGAGTAGTAGTGTCATTCTATCTGGGGGGTGGGGTGGGAGAGACAGCA 4008
QY 1001 agggggaagagtgaggaaacaataagcagcagtctgggagtcgggtgctatgctc 1060
DB 4009 AGGGGGAGGAGTTGGGAAGACAATAGCAGCATGCTGGGGATGCGGTGCGCTATGGAAC 4068
QY 1061 ctgagggggaaga----- 1074
DB 4069 CAGCTGGGCTCAGTCAGTGTATGACTGCGCGCGCATCCCTCGAGAGCTTGGCGTAA 4128
QY 1075 ----- 1074
DB 4129 TCATGTCATAGCTGTTCTCTGTGTGAATTTGTTATCCGCTCACAATTCCACACATA 4188
QY 1075 ----- 1074
DB 4189 CGAGCGGGAAGCATAAAGTGTAAAGCTGGGGTCTTAATGAGTGAAGCTAATCATTA 4248
QY 1075 ----- -accagctgcgcatlaa 1088
DB 4249 ATTGCGTGGCTCAGTCCCGCTTTCAGTCCGGGAACCTGTGCTGTCAGCTGCATTA 4308
QY 1089 tgaatcggaacagcgagggagagagcggttgcgtatctggcgctctccgctctcg 1148
DB 4309 TGAATGGGCCAAGCGCGGGGAGAGGGGTTTCCGTATTGGGCGCTCTTCCGCTCTCG 4368
QY 1149 ctactgactcgctcgctcgctcgttcgttcggtcgcgagcgagctatcactcaaaag 1208
DB 4369 CTCACTGACTCGCTGCGCTGCGTTCGCTGCGCTGCGGAGCGGATACGCTCACTCAAG 4428
QY 1209 ggggtataaggtatccacaagaatacagggatataagcgagaaagaacatgtagaacaa 1268
DB 4429 GCGGTATATAGGTTATCCACAGATCAGGGGATTAACGAGAAAGAACATGTGAGCAAAA 4488
QY 1269 ggcacaagaaggccaagaaacgtaaaagggcggttgcgttcgtttccctaagctc 1328
DB 4489 GGCACAGCAAAAGCCAGGAACCTTAATAAAGCGCGCTTGCCTGCTTTTCCATAGGCTC 4548
QY 1329 ggcgcctctgaagagatcaacaaatcgagctcaagtcagaagtggtgcgaacccgcaga 1388
DB 4549 CGCCCTCTGACGAGATCACAATAATCGACGCTCAAGTCAGAGGTGGCAAAACCCGACA 4608
QY 1389 ggaactaagaatacagaaggttcccccctggagctccctctgctgctctcctgttcg 1448
DB 4609 GGACTATAAAGATACAGAGGGGTTTCCCTCGAAGCTCCTCGTGGCTCTCCGTCCG 4668
QY 1449 accctgcgcttacgataaccgtctcgccttctcccttcgggaaacgctggcgcttct 1508
|||||

DB 4669 ACCCTGCCGCTTACCGGATACCTGTCCGCCCTTCTCCGGAAGCGTGGCCCTTCT 4728
QY 1509 caatgctacagctgtaagtatctcaattcagttcggtgaagtcgttcgctcaagctgagctc 1568
DB 4729 CAATGCTACAGCTGTAGTATCTAGTTCCGTGAGGTGCTTCCCTCAAGCTGGCTGT 4788
QY 1569 gtgcacgaaccccccttcacgcccagcagcgtctgctctatccgttaactatcgtctgag 1628
DB 4789 GTGCACGAACCCCCCGTTCAAGCCCGAGCCGCGGCTTATCCGTAATCGTATCGCTGAG 4848
QY 1629 tcaaacccggtlaagaacagactatcgccaactggaagcagccactggtlaaagattagc 1688
DB 4849 TCCAACCCGGTAAGACAGACTTATCCCACTGGCACAGCACTGTAACAGGATTAGC 4908
QY 1689 agagcgaggtatgtaggcggtgtcacagagttcttgaagtggtgcttaactagcttac 1748
DB 4909 AGAGCGAGTATGAGCGGGGCTACAGAGTTCTTGAAGTGCTGCTTAAGTACGAGCTAC 4968
QY 1749 actagaagagacagattgtgtatctgcctctgctgaagccagttaccttcggaacaaaga 1808
DB 4969 ACTGAAGAGACAGTATTTGGTATCTGGCTCTGCTGAAGCCAGTACCTTCGAAAAAGA 5028
QY 1809 gtgtgtagctcttgatccggcaacaaacacacgctgtgtagcgtgtgtttttgttgc 1868
DB 5029 GTTGTAAGCTCTTGTATCCGCAACAAACACACCGCTGTAGCGGTGTTTGTGTTGC 5088
QY 1869 aagcagcagattacgcgcagaaaaaagatctaaagaatcctttgacttttctacg 1928
DB 5089 AAGCAGCAGATTACGCCCAAAAAAGATCTCAAGAAATCTTTGATCTTTTACG 5148
QY 1929 gggctcagcctcagttgaaacgaacaaactcagttaaaggatttggctcagaagcgatc 1988
DB 5149 GGGTCTACCTCAGTGAAGCAAACTCACTTAAGGATTTGGTATGAGATTATCA 5208
QY 1989 aatttgatgattaga 2007
DB 5209 AAAAGATCTTCACTAGA 5227
|||

RESULT 13

US-08-940-371-49
Sequence 49, Application US/08940371
Patent No. 5851525
GENERAL INFORMATION:
APPLICANT: Ames, Robert S.
APPLICANT: Appelbaum, Edward R.
APPLICANT: Chaiken, Irwin M.
APPLICANT: Cook, Richard M.
APPLICANT: Gross, Mitchell S.
APPLICANT: Holmes, Stephen D.
APPLICANT: McMillan, Lynette J.
TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
TITLE OF INVENTION: Treatment of IL5 Mediated Disorders
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corp./Corporate
ADDRESSEE: Intellectual Property
STREET: P. O. Box 1539-UM220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/940,371
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 3, 2001, 15:21:54 ; Search time 5634.94 Seconds
(without alignments)
13613.261 Million cell updates/sec

Title: US-09-778-516-1

Perfect score: 8115

Sequence: 1 gatgtacggcgccagatatac.....gcgtgtcgtacgtacgagtc 8115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

```
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_est1a:*
34: em_est1b:*
35: em_est1c:*
36: em_est1d:*
37: em_est1e:*
38: em_est1f:*
39: em_est1g:*
40: em_est1h:*
41: em_est1i:*
42: em_est1j:*
43: em_est1k:*
44: em_est1l:*
45: em_est1m:*
46: em_est1n:*
47: em_est1o:*
48: em_est1p:*
49: em_est1q:*
50: em_est1r:*
51: em_est1s:*
52: em_est1t:*
53: em_est1u:*
54: em_est1v:*
55: em_est1w:*
56: em_est1x:*
57: em_est1y:*
58: em_est1z:*
59: em_est2a:*
60: em_est2b:*
61: em_est2c:*
62: em_est2d:*
63: em_est2e:*
64: em_est2f:*
65: em_est2g:*
66: em_est2h:*
67: em_est2i:*
68: em_est2j:*
69: em_est2k:*
70: em_est2l:*
71: em_est2m:*
72: em_est2n:*
73: em_est2o:*
74: em_est2p:*
75: em_est2q:*
76: em_est2r:*
77: em_est2s:*
78: em_est2t:*
79: em_est2u:*
80: em_est2v:*
81: em_est2w:*
82: em_est2x:*
83: em_est2y:*
84: em_est2z:*
85: em_est3a:*
86: em_est3b:*
87: em_est3c:*
88: em_est3d:*
89: em_est3e:*
90: em_est3f:*
91: em_est3g:*
92: em_est3h:*
93: em_est3i:*
94: em_est3j:*
95: em_est3k:*
96: em_est3l:*
97: em_est3m:*
98: em_est3n:*
99: em_est3o:*
100: em_est3p:*
101: em_est3q:*
102: em_est3r:*
103: em_est3s:*
104: em_est3t:*
105: em_est3u:*
106: em_est3v:*
107: em_est3w:*
108: em_est3x:*
109: em_est3y:*
110: em_est3z:*
111: em_est4a:*
112: em_est4b:*
113: em_est4c:*
114: em_est4d:*
115: em_est4e:*
116: em_est4f:*
```

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*

190: gb_est110:*
 191: gb_est111:*
 192: gb_hic:*
 193: em_gss_fun:*
 194: em_gss_hum1:*
 195: em_gss_hum2:*
 196: em_gss_hum3:*
 197: em_gss_hum4:*
 198: em_gss_hum5:*
 199: em_gss_hum6:*
 200: em_gss_hum7:*
 201: em_gss_hum8:*
 202: em_gss_hum9:*
 203: em_gss_inv1:*
 204: em_gss_inv2:*
 205: em_gss_inv3:*
 206: em_gss_other:*
 207: em_gss_pin1:*
 208: em_gss_pin2:*
 209: em_gss_pro:*
 210: em_gss_rod1:*
 211: em_gss_rod2:*
 212: em_gss_rod3:*
 213: em_gss_rod4:*
 214: em_gss_rod5:*
 215: em_gss_vit1:*
 216: em_gss_vit2:*
 217: em_gss_vit3:*
 218: gb_gss1:*
 219: gb_gss2:*
 220: gb_gss3:*
 221: gb_gss4:*
 222: gb_gss5:*
 223: gb_gss6:*
 224: gb_gss7:*
 225: gb_gss8:*
 226: gb_gss9:*
 227: gb_gss10:*
 228: gb_gss11:*
 229: gb_gss12:*
 230: gb_gss13:*
 231: gb_gss14:*
 232: gb_gss15:*
 233: gb_gss16:*
 234: gb_gss17:*
 235: gb_gss18:*
 236: gb_gss19:*
 237: gb_gss20:*
 238: gb_gss21:*
 239: gb_gss22:*
 240: gb_gss23:*
 241: gb_gss24:*
 242: gb_gss25:*
 243: gb_gss26:*
 244: gb_gss27:*
 245: gb_gss28:*
 246: gb_gss29:*
 247: gb_gss30:*
 248: gb_gss31:*
 249: gb_gss32:*
 250: gb_gss33:*
 251: gb_gss34:*
 252: em_gss_inv4:*
 253: em_gss_rod6:*
 254: em_gss_rod7:*
 255: em_gss_rod8:*
 256: gb_gss35:*
 257: gb_gss36:*
 258: gb_gss37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Db 732 CACCCGGTAAGACACGATTATCGCCACTGCGACGACCACTGCTAAGATTAGAG 791
Oy 1691 agcagatgtagtgcgcggtgtctacagatcttgaagtgtggtccctaactacggtacac 1750
|||||
Db 792 AGCAGAGTATAGCGGCTGCTACAGAGTCTTGAGATGGTGCGCTTAACCTACGCTAAC 851
Oy 1751 tagaagacagatattggtatctgcgtctgctgaagccaattacccctcgaaagaaagt 1810
|||||
Db 852 TAGAAGAACAGTATTGTTGATCTGCGCTCTGCTGAAGCAGTACTTCGGAAGAAAGT 911
Oy 1811 tggtagctcttgatccgcgcaacaacaccgcgtgtgtagcgt 1853
|||||
Db 912 TGGTAGCTCTTGATCCGGCAAAACACCGCTGTAGCGGT 954

RESULT 2

AU081137 1067 bp mRNA EST 15-NOV-1999
LOCUS AU081137 Oncorhynchus mykiss Kidney infected by infectious
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12,
mRNA sequence.
ACCESSION AU081137
VERSION AU081137.1 GI:6431485
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Proteanchtopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1067)
Sakai,M. and Kono,T.
The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus
JOURNAL Unpublished (1999)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: a0b208u@cc.miyazaki-u.ac.jp.

FEATURES
Source location/Qualifiers

1..1067
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="K12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"
/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"

BASE COUNT 235 a 292 c 288 g 250 t 2 others
ORIGIN

Query Match 9.5%; Score 772.6; DB 107; Length 1067;
Best Local Similarity 98.6%; Pred. No. 1.2e-204;
Matches 800; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

Oy 1076 caagctgcatatgaatgcgcaacgcgcggggaagcggttcgcatgtggcggtc 1135
|||||
Db 259 CCAAGTGCATTATATATGCGCCAAACGCCGGGAGAGCGGTTTCCGTAATGGCGCTC 318
Oy 1136 ttccgcttctcgtcactgaactgcgtgcgtcgttcgttcgctgcgcgcgcgtatc 1195
|||||
Db 319 TTCCGCTTCTGCTGACACAGCTCGGTGCGTGGCTGCTTCGGCTCGCGGAGCGGTATC 378
Oy 1196 agctactcaaaagcggttaacggttatccacaagaatacaggggataacgcaggaagaa 1255
|||||
Db 379 AGCTCACTCAAAAGCGGCTATACGGTTATCCACAGATCAGGGGATTAACGCAAGAAAGAA 438
Oy 1256 catgtagcaaaaagccagcaaaaagccaggaacggttaaaaagccgcgttctgtgcgtt 1315
|||||
Db 439 CATGTGAGCAAAAAGCGCACAAAAGCGCAGAAACGCTAATAAAGGCCGCGTGTGGCGTIV 498

Oy 1316 ttccatggtccgcgcccccctgacgacatcaaaaaatcgaagctcaagtcaagtgt 1375
|||||
Db 499 VTTCCATAGGCTCCGCCCCCTGACGAGCATCATAAAATGCACCTCAAGTCAAGAGGTG 558
Oy 1376 gcgaaaccggaagacacataaagaatacgaagcgtttcccccctggaagctccctgttcg 1435
|||||
Db 559 GGAATACCCGACAGGACTATTAAGATACCAAGGCTTCCCTGGAAGCTCCCTGCTGCG 618
Oy 1436 ctctccttccgaccctgcgcgccttacccgatatcctgtccgcttctcccttcgggaag 1495
|||||
Db 619 CTCTCCTGTCCGACCTGCGCCTTACCGGATACCTGTGTCGCTTTCCTCCCTCGGAAG 678
Oy 1496 cgtgcgcttctcctaagtcaagctgtagtatactcaagttcgtgtgtgttcgttc 1555
|||||
Db 679 CGTGGCGCTTCTCATTACCTACGCTGTAGGTATCTCAGTTGGTGTAGTTCGCTTC 738
Oy 1556 caagctggcgtgtgtgcagcaaaccccccggttaagccgaacccgtcgccttaacgta 1615
|||||
Db 739 CAGCTGGGCTGTGTGCAGAAACCCCGGTTGAGCCGACGCTCGCCTTATCCGGTAA 798
Oy 1616 ctatcgtctgagtcacaaccccgtaagaacagactatgcacactgcgcgcgcgcgcgt 1675
|||||
Db 799 CTATGCTGTGATGCTCAACCCGCTAAGACAGACTTATGCGCAGTGGCAGCCACTGG 858
Oy 1676 taacagatcagaagcagaggtatgtaggcggtcctacagagttcctgaagtgtggtc 1735
|||||
Db 859 TAACAGAGTATACAGACGCGATATGTAGCGGCTCTACAGAGTTCTTGAAGTGTGGCC 918
Oy 1736 taactaagctacactaagaagacagatattggtatctgcgtctgcgtcgaagcgttac 1795
|||||
Db 919 TAACTACGGCTACACTAGAGACAGATATTGCTATCTGCGCTGCTGCTAACCAGATTAC 978
Oy 1796 ctgcgaaaaagagttgtgtagctcttgatccgcgcaacaacaccgcgtgtgtagcggtg 1855
|||||
Db 979 CTTGGAATAAAGATTGGTAGCTTTGATCCGGCAAAACA-CAACGCTGTAGCGGT 1037
Oy 1856 ttttttgcgaagcagcagattacgcgc 1886
|||||
Db 1038 TTTTGTGTTTGC-AGCAGCAGATTACGCGC 1067

RESULT 3

AU081124 1089 bp mRNA EST 15-NOV-1999
LOCUS AU081124 Oncorhynchus mykiss Kidney infected by infections
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone KG'12,
mRNA sequence.
ACCESSION AU081124
VERSION AU081124.1 GI:6431472
KEYWORDS EST.
SOURCE rainbow trout.
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Proteanchtopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE 1 (bases 1 to 1089)
Sakai,M. and Kono,T.
The EST analysis of kidney and gill cells rainbow trout,
Oncorhynchus mykiss, infected with infectious hematopoietic
necrosis virus
JOURNAL Unpublished (1999)
COMMENT Contact: Masahiro Sakai
Faculty of Agriculture
Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Email: a0b208u@cc.miyazaki-u.ac.jp.

FEATURES
Source location/Qualifiers

1..1089
/organism="Oncorhynchus mykiss"
/db_xref="taxon:8022"
/clone="KG'12"
/clone_lib="Oncorhynchus mykiss Kidney infected by
infectious hematopoietic necrosis virus"

/tissue_type="Kidney infected by infectious hematopoietic
necrosis virus"
BASE COUNT 245 a 297 c 295 g 251 t 1 others
ORIGIN

Query Match 9.1%; Score 741.2; DB 107; Length 1089;

Best Local Similarity 96.5%; Pred. No. 7.5e-196;
Matches 801; Conservative 0; Mismatches 23; Indels 6; Gaps 4;

OY 1076 ccagctgcatatgaatcggccaacgacgaggaagagcggttgcgtatctggcgctc 1135
|||||
DB 260 CCACGTGATTAATGAATGCGCAACGGCGGAGAGCGGTGCTATTGGCGCTC 319
|||||
OY 1136 ttccgcttcctcgtcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1195
|||||
DB 320 TTCCGCTTCCTCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
|||||
OY 1196 agccactcaaaagcggttaaacggtatccacagaatcagggagatacagcagaagaa 1255
|||||
DB 380 AGCTCACTCAAAAGCGGTAATACGTTATCCACAGATCAGGGGTAACGACAGAAAGAA 439
|||||
OY 1256 catgtgcaaaaagcgacgcaaaagccaggaacggttaaaaagccgcttgcgtcgtc 1315
|||||
DB 440 CATGTGAGCAAAAGCGCAGCAAAAGCCAGAACCGTAAAAAGCCGCTGCTGCGCTT 499
|||||
OY 1316 ttccatagctccgccccctgacgagcatcacaataatcgacgtcctaagtcagaagt 1375
|||||
DB 500 TTTCCTAATAGCTCGCCGCCCTGACGAGCACAATAATGACGCTCAAGTCAGAGGTG 559
|||||
OY 1376 gcgaacccgcagagactataaagataccagcggttcccccgtgaagctccctcgtcgt 1435
|||||
DB 560 GCGAACCAGCAGAGACTATAAAGTACACAGGCGTTTCCCTGGAAGCTCCCTGCTGTCG 619
|||||
OY 1436 ctctcctgttccgacccctgcgtctaccggaatacctgtcgtccttccctcgtcgtgaag 1495
|||||
DB 620 CTCTCCTGTTCCGACCGCTGCGGCTTACCGGATACCTGTCGCCCTTCTCCCTGGGAG 679
|||||
OY 1496 cgtgagcttctcaatgctcagcgtgtagttatcagttcgtgtgtagtcgttcgtc 1555
|||||
DB 680 CGTGGCCCTTCTCATAGCTACGCTGATAGTATCTCAGTTGCTGAGTGTGCTGCTC 739
|||||
OY 1556 caagctgagctgtgtgacgagaccccccgttccagccgacgctgcgagctatccgtgaa 1615
|||||
DB 740 CAAGCTGGGCTGTGTGACAGAACCCCGCTTCAAGCCGCTGCGCTTATCCGGTAA 799
|||||
OY 1616 ctatcgtcttgaagtcacacccggtgaagacagactatcgcacttgcagcagcagcactg 1675
|||||
DB 800 CTATCGTCTTGAGTCAACCCGGTAAGACAGACTTATCCGCACTGGCAGCAGCACACG 859
|||||
OY 1676 taacaggaattagcagagcagagatgtatagcggtgtcacaagttcttgaag-tgtgtgc 1734
|||||
DB 860 TTAACGAGATTAGCAGACGAGAGTATGTAGGCGGTCTACAGATTCTTGAAGTTGTGTCG 919
|||||
OY 1735 ctactcagctgactaactaagaagacagatattgttatctgctgctgctgtaagcaacta 1794
|||||
DB 920 CTAATCTACGGCTTACACTAAGAAAGAACTATTGTATCTGCTGCTGCTGAGCCAGTTA 979
|||||
OY 1795 ccttcgcaaaaagagttgttagctctt-gatccgcaaaaacaacaa--ccgctgtgtagc 1850
|||||
DB 980 CCTTCGGAAAAAGAGTGGGTAGCTCTTGAGATCCGGCAAAACACCCGTGGGAAGG 1039
|||||
OY 1851 ggtgtgttttttctgttgcagca-gcagattacgcgcaaaaagagat 1899
|||||
DB 1040 GGGGTCTTTTGTTCACACGACGAGATATCCCGGAAAAAACCGGAT 1089
|||||

RESULT 4
AJ281449 800 bp mRNA EST 30-JUN-2000
LOCUS AJ281449 800 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION AJ281449

VERSION AJ281449.1 GI:6929329
KEYWORDS EST
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
1 (bases 1 to 800)
REFERENCE Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M. B.
and Kafatos, F. C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
JOURNAL Contact: Dimopoulos G
MEDLINE Foltis C. Kafatos Laboratory
COMMENT European Molecular Biology Laboratory
Meyeroofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. 800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/clone_id="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: EcoRI; Site 2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN

Query Match 8.7%; Score 708; DB 104; Length 800;
Best Local Similarity 96.7%; Pred. No. 1.3e-186;
Matches 723; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 1260 tgaagcaaaaagcgacgcaaaaagcggaacggttaaaaagcggttgcgttctc 1319
|||||
DB 12 TAAAGTTCGGGCGCAAGCAAAAGCCAGAACCGTAAAAAGCGGCTTGCCTTTTC 71
|||||
OY 1320 cataggtccgccccctgacgagcatcacaataatcgacgtcctaagtcagaagtgcga 1379
|||||
DB 72 CATAGGCTCCGCCCCCTCGAGAGCATCATCAAAATAGACCTCATAGCAAGGTGGCGGA 131
|||||
OY 1380 aaccgacagagactataaagataccagcggttcccccgtgaagctccctcgttcgctc 1439
|||||
DB 132 AACCGAGAGACTATAAGATACAGGCGTTTCCCTGGAAGCTCCCTGCTGCTCT 191
|||||
OY 1440 cctgttccgacccctgcgttaccggaatcctgtccgcttccctcgttcggaagcgtg 1499
|||||
DB 192 CCGTGTGCGACCGCGGCTTACCGGATACCTGTCGCCCTTTCCTTCCGGGAAGCGTG 251
|||||
OY 1500 ggcgttctcaatgctcagctgtatgataatcagttcgtgtgtatggttcgttcgaag 1559
|||||
DB 252 GCGGTTTCTCATATACCTACGCTGATGATTCAGTTGCGGTGAGTGTGCTTCACAG 311
|||||
OY 1560 ctggagctgtgtgacggaaccccccgttaccgacgacgctgcgcttaccgtaactat 1619
|||||
DB 312 CTGGGCTGTGTGCAAGAACCCCGCTTACGCCGACGCTCGCTTATCCGGTAACAT 371
|||||
OY 1620 cgttctgagtcacacccggttaagaacagactatcgcacttgcagcagcactgtgtaac 1679
|||||
DB 372 CGTCTTGATGTCACACCGGTAAGACAGACTTATTCGCCACTGCGAGGACCACTGTAAC 431
|||||

QY 1680 aggattacagagcaggtatgtaagcgtgtctacagattcttgaatgtgtgcctaac 1739
|||||
Db 432 AGGATTACAGAGCAGGATATGTAGCGCGTGTCTACAGATTCTTAAAGTGGGCTAAC 491
QY 1740 tacggtcacactagaagacagatttggtaicctgcgtctgcgtgaagccagttacctc 1799
|||||
Db 492 TACGGCTACACTAGAAAGACAGATTATTTGTATCTCGCTCTGCTGAAGCCATTACCTTC 551
QY 1800 ggaataaagtgtgattgtcttgcgtccgcaaacacacacgcgtggtgaagtggtttt 1859
|||||
Db 552 GGAATAAAGATTGGTAGCTCTTGATCCGGCAACACACCGCTGCTGATGCGGTGTTT 611
QY 1860 ttgtgtcagacagcagattacgcgcagaaataaaggatctcaagaagatccttgatc 1919
|||||
Db 612 TTTGTTTCAAGCAGCAGATTATACGGCAGAAAGATCTCAAGATCTTGTGATC 671
QY 1920 ttctcaggggtctgcagcctcagtggaagaaatacagttcaaggatttggatcatg 1979
|||||
Db 672 TTTTCTACGGGCTCTGACGCTCAGTGAACGAAACCTCAGTTAAGGATTTTGGTCAAG 731
QY 1980 agcggtacatattgtaattgattga 2007
|||||
Db 732 AGATTATCAAAAAGATCTTCACTTAGA 759

RESULT 5

AJ281480 1004 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION AJ281480
VERSION AJ281480.1 GI:5929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anopheles.
1 (bases 1 to 1004)

REFERENCE

AUTHORS Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Anserge, W., Soares, M. B.
and Kafatos, F. C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

JOURNAL

COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
MEDLINE 20300950
SOURCE Location/Qualifiers

FEATURES

1..1004
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/cell_line="Anopheles gambiae immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match

8.6%, Score 700.6; DB 104; Length 1004;

Best Local Similarity 97.3%; Pred. No. 1.7e-184;
Matches 712; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1276 aaaagccaggaacgcgttaaaaagccgcgttgcgttcttccataggtccgcgc 1335
|||||
Db 1 AAAAGCCAGGAAACCGTTAAAGAGCCGGCTGCTGCGCTTTTTCATAGCTCCGCC 60
QY 1336 ctgacagcatcacaaaaatcacgcctcaagtcagagtggtgcgaaccagagactat 1395
|||||
Db 61 CTGACGAGCATACAAAAATCAGCCTCAAGTCAGAGTGGGGAACCCGACGACATAT 120
QY 1396 aaagaatacagagcgttcccccctggaagtcctcgtgcgtctcctgtccgaacctgc 1455
|||||
Db 121 AAAGTACACAGCGCTTCCCTCGAAGCTCCCTGCGGCTCTCTGTTCCGACCTGCG 180
QY 1456 cgttacccggaactgtgcgcgttccccccttcccttgggaagcgtgcgtcttccaaagt 1515
|||||
Db 181 CGCTTACCGGATACCTGTCGCCCTTCTCCCTTGGGAAGCGTGGCGCTTCTCATAGCT 240
QY 1516 cagcgttagatcatcagttcagtcggtcgtcgttcgcctcgaagctggtgtgcag 1575
|||||
Db 241 CACGCTGATAGTATCTCAAGTTGGTGTAGTGTGCTGCCAAGCTGGGCTGTNCGACG 300
QY 1576 aacccccgttccagccagccgcgtgcgtctccgttaactatcgtcttgcctcaacc 1635
|||||
Db 301 AACCCCGCTTCAGCCCGCACCGCTGCTTATCCGTTAATATGCTTGTGATCCAAAC 360
QY 1636 cgttaagcagactatctgcgcacgtgcgcagcgcacgtgtaacagattagcagaagcga 1695
|||||
Db 361 CGGTAAAGACACGATTTATGCGCCTGACGACGACGCTGTAAAGATTACAGAGCGA 420
QY 1696 ggtatgtagcgtgtcctacagagttcttgaagtggtgccttaactacagcgtctacatagaa 1755
|||||
Db 421 GGTATGTAGCGCGTCTACAGAGTTCTTGAAGTGTGGCTTAACATGAGCTACACTAGAA 480
QY 1756 ggaacgatttggatctgcgtcgtctgcgtgaagccagttaccttcggaaaagattgta 1815
|||||
Db 481 GAACAGTATTTGTGATCTGCGCTGCTGTAACCGATTAACCTTCGGAAGAGTTGGTA 540
QY 1816 gctcttgatccggaacaaacacacgcgtgtagagtggttttttttttgcgaagcagc 1875
|||||
Db 541 GCTCTTGATCCGCGAACAACACCGCTGCTGAGGGGTGTTTGTGTTGCAAGCAC 600
QY 1876 agattacgcgcaaaaagagatctcaagaagatccttgccttcttctacaggggtctg 1935
|||||
Db 601 AGATTACGCGCAGAAAAAGATCTCAAGAGATCTTGTGATCTTTTCTACGGGGTCTG 660
QY 1936 acgttcagtggaagaaatctcaagtltaaggatttggtaagcgagatacatattg 1995
|||||
Db 661 ACGCTCAGTGAACCAAACTCAGCTTAAGGATTTGTGATGAGATTATCAAAAAAGGA 720
QY 1996 aatgtattaga 2007
|||||
Db 721 TCTTCACTTAGA 732

RESULT 6

AJ281616 700 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P8A12-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P8A12, mRNA sequence.
ACCESSION AJ281616
VERSION AJ281616.1 GI:6929495
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anopheles.
1 (bases 1 to 700)
REFERENCE Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
AUTHORS Donohue, M., Schultz, J., Benes, V., Bork, P., Anserge, W., Soares, M. B.
and Kafatos, F. C.

TITLE	Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)					
MEDLINE	20300950					
COMMENT	Contract: Dimopoulos G Fotis C. Kafatos Laboratory European Molecular Biology Laboratory Meyerhofstrasse 1, 69117 Heidelberg, Germany. Location/Qualifiers 1..700 /organism="Anopheles gambiae" /strain="4A r/f" /db_xref="taxon:7165" /clone="4A3A-P8A12" /clone.lib="Anopheles gambiae immune competent 4A3A" /cell.line="immune competent 4A3A" /lab.host="E. coli DH10B" /note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."					
FEATURES	source					
BASE COUNT	167 a	196 c	189 g	146 t		
ORIGIN						
Query Match	8.3%	Score 677.4;	DB 104;	Length 700;		
Best Local Similarity	98.8%;	Pred. No. 4.8e-178;				
Matches 681;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;		
Oy 1163	ggcgccgcgfcgttgctgctcgtgcgcgcgcgcgatcacgaactcaagaagcggatatacggtt	1222				
Db 12	GGTTGCGGTCTTCCGCCTCAGCGGGAACCATCACTCAAAGCGGTAATACGGTTT	71				
Oy . 1223	atccacaagaatcaaggagaaacacgaagaaaacattgaccaaaagccagcaaagtc	1282				
Db 72	ATCCACAGAATACAGGGGAAACGCCAGAAAACATGTACGAAAAAGCCACAAGACC	131				
Oy 1283	caggaacccgtaaaaagcccgcgtctgcgttgttttccaiaagctccgcccttcagca	1342				
Db 132	CAGGAACCGTAAAAAGCCGCGCTTGCTGCGGNNTTCATAAGCTCCGCCCCCTGACGA	191				
Oy 1343	gatatacaaaatcgaacctcaagtcatgacagagctggcgaaaaccagacagatatagaata	1402				
Db 192	GCATACACAAAAATCACACTCAAGTAGAGGAGGGGAAACCCGACAGACTATTAAAGATA	251				
Oy 1403	ccaggcgtttccccctctggaagctccactctgtgcgtcttccctgttcgaacctgcagttac	1462				
Db 252	CCAAGCGTTTTCCCCCTGGAAAGCTCCCCTCGGCGCTCTCTGTTCCGACCCCTGCCATTAC	311				
Oy 1463	cggataacctgcgcctttctcccttcgsgaaagcgtggcgcgttttccaigtccagctg	1522				
Db 312	CGGATACCTGTCGCGCTTCTCCTTCGGGAAGCGTGCGGCTTCTCATATAGCTACGGCTG	371				
Oy 1523	tagtatcatctcagttcgtggttagtgctgcgtccaaagctggcgtgtgtgcacgaacccc	1582				
Db 372	TAGGATATCTCAATTTGGGTGTAGGTGGTTCCGCTCCAAAGTGGGCTGTGTGCAGAACCCCC	431				
Oy 1583	ggttaagcccccagcgtgcgaccttacgcgtaacatagttcttggttccaacccggttag	1642				
Db 432	CGTTACGCCGACCGCTGGCCCTTATCCGGTAACATGATGCTCTTGATCCAACCCCGTAAG	491				
Oy 1643	aacagacctatgcagctatgcagcagccacactgtttaeaggatttaacagacgagttatgt	1702				
Db 492	AACAGCATTATGCCACACTGGCAGCAGCCACTGTGTAAACAGATTATACAGACGACGATATGT	551				
Oy 1703	aggcgggtctacagaagttctvtgaagtgtgtgacctaatcaggtctaactagaagacagt	1762				
Db 552	AGCGCGTGTATACAGAGTTCTTGAAGTGTGGGCTTAATACAGGCTACACTAGAAAGAAGT	611				

QY	1763	atgtgcatctcgccttgcctaaccagttaccccttggaataaaagtggagctctg	1822
Db	612	ATTGTGATCTTCGGCTCTCCTAGACCAATTAACCTTCGAAAAAGAATTGGACTTGG	671
QY	1823	atccggcaacaaccacccgcgtcgtagcg	1851
Db	672	ATCCGCCAACAAACCACCGCTGGTAGCG	700
RESULT	7		
AJ281552			
LOCUS			
DEFINITION	AJ281552	1070 bp	mRNA EST 30-JUN-2000
ACCESSION	4A3A-P6F11-F	Anopheles gambiae immune competent 4A3A Anopheles	
VERSION	gambiae cDNA clone 4A3A-P6F11,	mRNA sequence.	
KEYWORDS	AJ281552.1	GI:6929432	
SOURCE	EST.		
ORGANISM	African malaria mosquito.		
REFERENCE	Anopheles gambiae		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
	Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae		
	; Anophelinae.		
	1 (bases 1 to 1070)		
	Dimpoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,		
	Donohue,M., Schultz,J.J., Benes,V., Bork,P., Ansoorge,W., Soares,M.B.		
	and Katatos,F.C.		
TITLE	Anopheles gambiae pilot gene discovery project: identification of		
JOURNAL	mosquito innate immunity genes from expressed sequence tags		
MEDLINE	generated from immune-competent cell lines		
COMMENT	Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)		
	Contact: Dimpoulos G		
	Fotis C. Katatos Laboratory		
	European Molecular Biology Laboratory		
	Meyerhofstrasse 1, 69117 Heidelberg, Germany.		
FEATURES			
Source	Location/Qualifiers		
	1..1070		
	/organism="Anopheles gambiae"		
	/strain="4A r/f"		
	/db_xref="taxon:7165"		
	/clone="4A3A-P6F11"		
	/clone.lib="Anopheles gambiae immune competent 4A3A"		
	/cell_line="immune competent 4A3A"		
	/lab_host="E. coli DH10B"		
	/note="Vector: pUT7SD-Pac (Pharmacia) with a modified		
	polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from		
	forward priming site which reads from the 3' end of the		
	cDNA. The 4A3A is a directionally cloned and normalized		
	cDNA library that was constructed from the 4A3A cell line		
	oligo-T primed cDNA according to: Bonaldo, Lennon & Soares		
	(1996) : Normalization and Subtraction: Two Approaches to		
	Facilitate Gene Discovery, Genome Research 6, 791-806."		
BASE COUNT	263 a	283 c	255 g
ORIGIN	269 t		
Query Match	8.3%;	Score 672.6;	DB 104; Length 1070;
Best Local Similarity	97.3%;	Pred. No. 1.2e-176;	
Matches 684; Conservative	0;	Mismatches 19;	Indels 0; Gaps 0;
QY	1305	ttagcggcgatttttcataagcttcgcccccctgcgcgagcatcacaaaaatgcaggcca	1364
Db	1	TTTGCTGGCGCTTTTCATATAGGCTCCGCCCCCTGCAGCATCACAAAATGAGCGCTCA	60
QY	1365	agtcagagggtgcgaaaaccgcagcagactataaataccaaagcgcttcgccctggaagc	1424
Db	61	AGTCAGAGGTGCGGAACCCGACGAGACTATAAAGATACAGGCGCTTCCCCTGGAAGC	120
QY	1425	tcacctgctgagctctcctgcttcagaacctgcgcgttacccgatcctgtccgcctttcc	1484
Db	121	TCCCTCGAGCGCTCTCCTGTTCCGACCCCTGCAGCTTACCGGATACCTGTCGCCCTTTCTC	180

FEATURES	COMMENT	ORGANISM	VERSION	KEYWORDS	ACCESSION	DEFINITION	LOCUS	RESULT
Db	181	CTTTCGGGAAGGAGTGGCCCTTTCTATATGCTACACGCTATAGGATCTCATGTTGGGTAG	240					1545
Db	241	GTCCGTTCCGCTCCAACTGGCGTGTGTGACACGACACCCCGTTCAAGCCGACCGCTGGCC	300					1544
Db	1605	ttatccggttaactatcgctcttgatgcacaaccgcgttaagaacagactatcgccatggca	1664					1544
Db	301	TTATCCGGTAACACTACGCTTGTAGTCACACCCGGTAAACACGACTATCGCCACTGGCA	360					1544
Db	1665	gaagccactgtgaaacagattagcagcagcagcagcagcagcagcagcagcagcagcagc	1724					1544
Db	361	GCAGCCACTGGTAAACAGGATTTAGCAGACGAGAGTATGAGGGGGCTTACAGAGTTCTTG	420					1544
Db	1725	aagttgtgccaactacacgctacacactagaagacagatatttgatctgcgctctg	1784					1544
Db	421	AGTGTTGGCCCTCAACTACGGCTACACCTAGAGAAACAGATTGTTGATCTGGCGCTGTG	480					1544
Db	1785	aagccagttaccccttcggaaaaagattgtgagctcttgatcccgcaacaaccacgcct	1844					1544
Db	481	AAGCCAGTTACTCTCGGAAAAAGATTGGTATGCTTGTATCCGGCAACCAACCCGCT	540					1544
Db	1845	ggttagcgtgtgttttttggttgcaagcagcagatatacggcagagaaaaaagatctaa	1904					1544
Db	541	GGTACCGGGGTTTTTTTTTGTGCAAGGACGAGATTACCGCGGAGAAAAAGAGACTTAA	600					1544
Db	1905	gaagatcccttgcattcttcttcagcggcgtcgcctcagtgtaagcaaaactcagctaa	1964					1544
Db	601	GAAAGTCTCTTTCATCTTTTCTTCTACGGGGCTGTACGCTCAGTGGAAAGAAAACTCACGTTAA	660					1544
Db	1965	gggatttggatcagcggatcacatatttgaatgattaga	2007					1544
Db	661	GGGATTTGGTCATGAGATTATCAAAAAGAGACTTCACCTAGA	703					1544
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
MEDLINE								
COMMENT								
FEATURES								
source								

polylinker: site:1: EcoRI: site:2: NotI: sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line o1igo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806.."					
BASE COUNT	159 a	204 c	193 g	147 t	
ORIGIN					
Query Match	8.2%	Score 663;	DB 104;	Length 703;	
Best Local Similarity	97.8%;	Pred. NO. 5.2e-174;			
Matches 672; Conservative	15;	Mismatches	0;	Gaps	0;

[illegible]

ORGANISM	Chlamydomonas reinhardtii								
REFERENCE	Eukaryota,Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;								
AUTHORS	Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C., 1 (bases 1 to 757)								
TITLE	Lefebvre,P., McDermott,J.P., Shreger,J., Sillflow,C. and Sten,D. Analyses of the Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3 unpublished (2000)								
JOURNAL	Contact: Charles Hauser								
COMMENT	DCMB Box 91000 Duke University Durham, NC 27708-1000 Tel.: 919 613 8159 Fax: 919 613 8177 Email: chausser@duke.edu.								
FEATURES									
SOURCE	location/Qualifiers 1..757 /organism="Chlamydomonas reinhardtii" /strain="CC-1690 wild type mt+ 21gr" /db_xref="taxon:3055" /clone_lib="C. reinhardtii CC-1690, stress condition I, normalized, lambda zap II" note="Vector: pBluescript II SK-; site.1: EcoRI; site.2:- XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min, 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr , 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda zap clones by superinfection with Exsistit (Stratagene) phase. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."								
BASE COUNT	167 a	219 c	208 g	161 t	2 others				
ORIGIN									
Query Match	8.0%; Score 653.2; DB 170; Length 757;								
Best Local Similarity	95.3%; Pred. NO. 3e-171;								
Matches	684;	Conservative	0;	Mismatches	33;	Indels	1;	Gaps	1;
Dn	1035	tggggaacgcgttgyggctcattgcttccttcgaagcggaagaacaagctgatcaataagaaac	1094						
Dd	34	tgggtaaccggcccccgcgtttccgttcgggaaaacctgctgcacactgcattaatgaatnc	93						
Oy	1095	ggccaagcgcggggaaggagggcggttgcatgttggcgctcttcgcgtctctctgctaact	1154						
Dd	94	ggccaacggcgccggggaagggcggtttggcttatggggccctcttcggcttccgtctcact	153						
Oy	1155	gaactcgctgcgctgctgctcgttcgtcgtctgctggcgagcgatcacgtcattcaaagcgcta	1214						
Dd	154	gactgcgtgcgcgtcgttcgcttgcctgcggcgagcgagctatcagctcaatccaaagcgcta	213						
Oy	1215	atacggtattcacagaatlcaggggataacgcgaagaaagaaacatgtgtgacaaagccag	1274						
Dd	214	ataccggttatccacagaaatccagggatgaacgcagaaacacatgatgacaaaaagccag	273						
Oy	1275	caaaagcgcaagaaacgttaaaaagccgcgttgtgcgtggcggtttttccataagctccgcccc	1334						
Dd	274	caaaaagcccgagaaacgttaaaaagccgcgttgtgcgtggcggtttttccataagctccgcccc	333						
Oy	1335	cctgcagagcaatcacaaaatcgacgtcacaagtcaagagtgtagcgaaaccgcagagacta	1394						
Dd	334	cctgcagagcaatcacaaaatcgacgtcacaagtcaagagtgtagcgaaaccgcagagacta	393						
Oy	1395	taaagataccaggcggttcccccttggaaagctccctcgtgcgctcbctgtttccagacctg	1454						
Dd	394	taaaatattccagagcctttttcccttggaagctccctcgtgcgctcbctgtttccagacctg	453						

QY	1455	ccgttaccggtatacctgttcgcgccttcttcctcttgggaaagctgtgaccttcttaatgc	1514
Db	454	CCGCTTACCGGAAACCTGTGTCGGCTTTTCCTTCGGGAAGCGTGGCTTTCATAGC	513
QY	1515	tcacgtctgaagtatctcaagttcgtgtgtaagtcgtcttcgcgtcccaagctgagctgtgtgac	1574
Db	514	TCAAGCTGTAGGATATTCAGTTGCGTGTAGTGTTCCTCCAGACTGGGCTGTGTGCAC	573
QY	1575	gaacccccgttcaagccgcagcgcgtgagccttaccgtaactatcgtctttagtccaac	1634
Db	574	GAACCCCCGTTACAGCCCAACCGCTCGCCTTATCCGGTAACTATGTTCTTGAGTCCAAC	633
QY	1635	ccgtgtaaagcacagacttatccgcacatgagcagcagccacatgataaggttaagcaagc	1694
Db	634	CCGGTAAAGACAGACTTATTCGCACATGGCAGCGCCACTGTGTAAACAGGATTTACAGAGCG	693
QY	1695	aggtatgtatggcgtgtctctcaagatctcttgaagtgtgtgcctaactacagctcacta	1752
Db	694	AGGTATGTATGGCGGTCTCTACAGAGTTCTTGTGAGTGAAGTACTAAC-ACGGCTACACTA	750
RESULT	10		
LOCUS	AL039459		
DEFINITION	DKFZP43401210.s1.434 (synonym: htes3) Homo sapiens cDNA clone		
ACCESSION	AL039459		
KEYWORDS	AL039459.2 GI:5935158		
SOURCE	EST.		
ORGANISM	human.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 734)		
AUTHORS	Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann		
	,S.		
	EST (Duesterhoeft, et al.)		
	Unpublished (1999)		
	On Jul 7, 1999 this sequence version replaced gi:5408507.		
COMMENT	Contact: Duesterhoeft A		
	MIPS		
	Am Klopferspitz 18a D-82152 Martinsried, Germany		
	This is the 3' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer		
	Research Center (DKFZ); Email:s.wiemann@dkfz-heidelberg.de;		
	sequenced by Jagen (Hilden/Germany) within the cDNA sequencing		
	consortium of the German Genome Project.		
	No r1 sequence available.		
	This clone (DKFZP43401210) is available at the RZPD in Berlin.		
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059		
	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..734		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="DKFZP43401210"		
	/clone_1lb="434 (synonym: htes3)"		
	/tissue_type="testis"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"		
BASE COUNT	172 a 199 c 194 g 161 t 8 others		
ORIGIN			
Query Match	7.7%; Score 621.4; DB 105; Length 734;		
Best Local Similarity	94.4%; Pred. No. 2.4e-162;		
Matches	685; Conservative 0; Mismatches 33; Indels 8; Gaps 4		
QY	1159	cgctgcgcctcgcgtcttcgcgtcgcgcgcgcgtatcagctcaactcaaaagcgcgtaatac	1218
Db	9	CGCTGCGCGCTCGCTGTCGCGTGCAGCGCGTATCACCTCACTCAAAAGCGCGTAAATAC	68

QY 1496 cgtgagcttctcaatgctcagctgtaggtatctcagttcggtgtagtgcgtc 1555
|||||
Db 964 CGTGGCGCTTTCATACCTACCGCTGAGGATCTCAGTTGCGTAGTCGCTC 1023
QY 1556 caactggcgctgtgacgaacccccggttcagccgacgctgagctatccggtaa 1615
|||||
Db 1024 CAACCTGGCGCTGTGCGAGAA-CCCCGTTACGCCACCGCTGCGTTATCCGGTAA 1082
QY 1616 ctatcgcttgaatcccaacccggttaaacacagactatcgcgaactgacagacactgg 1675
|||||
Db 1083 CTATCGCTTTCATACCGCTGAGGATCTCAGTTGCGTAGTCGCTC 1142
QY 1676 taacagattagcagcagcag 1696
|||||
Db 1143 TAACAGATTAGCAGAGCGAG 1163

RESULT 14
AJ281699 636 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P9E3-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P9E3, mRNA sequence.
ACCESSION AJ281699
VERSION AJ281699.1 GI:6929578
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anopheles.
1 (bases 1 to 636)
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultze,J., Benes,V., Bork,P., Ansong,W., Soares,M.B.
and Kafatos,F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .636
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P9E3"
/clone_1id="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-r primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 152 a 179 c 166 g 139 t
ORIGIN

Query Match 7.4%; Score 604; DB 104; Length 636;
Best Local Similarity 98.9%; Pred. No. 1,76-157;
Matches 629; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1202 ctcaaaagcgtaatacgtatcacaggaatcagggatacaggaagaagaacatg 1261
|||||
Db 1 CTCAAAAGCGGTAATACCGTTATCCACAGATCAGGGGATACCGAAGAAACATGTG 60

QY 1262 agcaaaagccagcaaaagccaggaacgtaaaagccggttgcgttttcca 1321

Db 61 AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGCCCGGTTGCGCTTTTTC 120
QY 1322 taggtccgccccccctggcagagcatcacaaaaatcagaactaaagttaggtggcga 1381
Db 121 TAGGCTCGCCCCCTTGAGAGCATACAAAATGACGCTTAATGACAGAGTGCGA 180
QY 1382 cccagcagaactataaataacacagcgcttcccccctggaagctccctgtgctctcc 1441
Db 181 CCGGACAGAGCATATAATACATACAGCGTTTCCCTGGAAGCTCCCTGTCCTCC 240
QY 1442 tgttcgaacctgcccgttaacggaactcgtgcgcttctccctcgggaagcgtgac 1501
Db 241 TGTTCGACCCCTGCGCTTACCGGATACCTGCCCCCTTTCCTTCGCGGAAGCTGCG 300
QY 1502 gcttctcaatgctcagctgtaggtatcgaatgctggtgtaggtcgttgcctcaagct 1561
Db 301 GCTTCTCATAGCTACCGCTGAGGATCTCAGTTGCGGTGATGCTTCCTCCAGCT 360
QY 1562 gggctgtgtgcaagaccgccggttcagccgacccgctgcgctatccgtaactatcg 1621
Db 361 GGGCTGTGTGCAAGAACCCCGCTTACAGCCGACCGCGCTTATCGGTAACATCG 420
QY 1622 tcttgatccaaacccggttaagacacgactatcgcactgycagcagcactggttaacag 1681
Db 421 TCTTGATCTCAACCCCGGTAGACACGACTTATCCGCTGGAGAGCCACTGTGTAACAG 480
QY 1682 gattagcagaacgaagttgttaggcgtgtgtacagaagttcgttaagttgtggtc-taac 1740
Db 481 GATTAGCAGAGGAGGATGTATGCGGCTGTACAGATTTCTTGAAAGTGTGCGCTTTTCT 540
QY 1741 acggttaactagaagagcagatattgttatctgcgctctgtgaagcagttacttcg 1800
Db 541 ACGGCTACACTAGAGAACAGATTTGGATGCTGGCTGCTGTAAGCAGTTACTTCG 600
QY 1801 g-aaaaagattggtagctcttgatccggcaacaa 1835
Db 601 GAAAAAAGATTGTTAGCTTTGATCGCGCAACAA 636

RESULT 15
AO876119/c 784 bp DNA GSS 08-NOV-1999
LOCUS AO876119
DEFINITION V133E1 mtn-3xHA/LacZ insertion library, strain Y2278 Saccharomyces
cerevisiae genomic 5', DNA sequence.
ACCESSION AO876119
VERSION AO876119.1 GI:6288363
KEYWORDS GSS.
SOURCE Saccharomyces cerevisiae.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetaceae; Saccharomycetaceae; Saccharomyces.
1 (bases 1 to 784)
Ross-Macdonald,P., Roemer,T., Coelho,P.S.R., Agarwal,S., Kumar,A.,
desBrelvi,S.A., Cheung,K.-H., Sheehan,A., Symoniatas,D., Jansen,R.,
Umastley,L., Heldman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished (1999)
JOURNAL Contact: Kumar A
COMMENT Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mtn-3xHA/LacZ insertion.
Seq primer: GGCCTCTTCTTGTGGAAGTAC
Class: transposon-tagged.
Location/Qualifiers
1. .784
/organism="Saccharomyces cerevisiae"

THIS PAGE BLANK (USPTO)